

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:09:35 : Search time 22 seconds
(without alignments)
5988.902 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 4120

Sequence: 1 ggaagtgaagtagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-DB=Issued_Patents_AA -OFMT=fastap -SUFFIX=n2p.ra -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -DUALIGN=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DUALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170 -CGN_1_1_17 -runat_04042003_090918_20329 -NCPU=6 -ICPU=3
-NO_XLPAY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1156.5	28.1	515	2	US-09-073-362-3
2	1156.5	28.1	515	2	US-09-243-920-3
3	1131.5	27.5	511	2	US-09-073-362-1
4	1131.5	27.5	511	2	US-09-243-920-1
5	621	15.1	241	2	US-08-825-781-3
6	615	14.9	241	2	US-08-825-781-4
7	524.5	12.7	245	2	US-08-825-781-1
8	286	6.9	629	2	US-08-132-990A-8
9	286	6.9	629	5	PCT-US92-09382-8
10	282.5	6.9	629	5	US-09-134-001C-4290
11	279.5	6.8	622	2	US-09-132-990A-4
12	279.5	6.8	622	5	PCT-US92-09382-4
					Sequence 3, Appli
					Sequence 3, Appli
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 4, Appli
					Sequence 1, Appli
					Sequence 8, Appli
					Sequence 4290, Ap
					Sequence 4, Appli
					Sequence 4, Appli

13	224.5	5.4	499	4	US-09-134-001C-5370	Sequence 5370, Ap
14	200.5	4.9	494	4	US-09-134-001C-4475	Sequence 4475, Ap
15	196	4.8	476	4	US-09-134-001C-3778	Sequence 3778, Ap
16	194	4.7	465	4	US-09-134-001C-4428	Sequence 4428, Ap
17	192.5	4.7	367	2	US-08-132-990A-2	Sequence 2, Appli
18	192.5	4.7	367	5	PCT-US92-09382-2	Sequence 2, Appli
19	190.5	4.6	456	4	US-09-134-001C-3771	Sequence 3771, Ap
20	178	4.3	344	4	US-09-134-001C-4175	Sequence 4175, Ap
21	169.5	4.1	503	4	US-09-068-195-24	Sequence 24, Appli
22	166	4.0	453	2	US-08-132-990A-6	Sequence 6, Appli
23	166	4.0	453	5	PCT-US92-09382-6	Sequence 6, Appli
24	163.5	4.0	482	4	US-09-134-001C-4309	Sequence 4309, Ap
25	158.5	3.8	414	4	US-09-134-001C-3357	Sequence 3357, Ap
26	158	3.8	496	4	US-09-134-001C-3001	Sequence 3001, Ap
27	141	3.4	1212	4	US-09-268-866-2	Sequence 2, Appli
28	130.5	3.2	165	4	US-08-858-207A-462	Sequence 462, App
29	129	3.1	249	4	US-09-134-001C-5030	Sequence 5030, Ap
30	129	3.1	430	4	US-09-134-001C-2981	Sequence 2981, Ap
31	126	3.1	454	4	US-09-134-001C-3501	Sequence 3501, Ap
32	123.5	3.0	432	4	US-09-370-253-12	Sequence 12, Appli
33	123.5	3.0	527	4	US-09-134-001C-3358	Sequence 3358, Ap
34	119.5	2.9	3177	2	US-08-477-451-4	Sequence 4519, Ap
35	117	2.8	349	4	US-09-134-001C-4519	Sequence 5109, Ap
36	115.5	2.8	554	4	US-09-134-001C-5109	Sequence 3702, Ap
37	112	2.7	473	4	US-09-637-118B-2	Sequence 8, Appli
38	111.5	2.7	413	4	US-09-134-001C-3702	Sequence 10, Appli
39	109	2.6	525	4	US-09-351-224E-8	Sequence 11, Appli
40	108	2.6	599	1	US-07-879-617A-10	Sequence 4, Appli
41	108	2.6	599	1	US-08-295-814E-11	Sequence 10, Appli
42	108	2.6	599	1	US-08-240-783B-4	Sequence 4, Appli
43	108	2.6	599	1	US-08-753-985-10	Sequence 10, Appli
44	108	2.6	599	3	US-09-084-813-4	Sequence 4, Appli
45	108	2.6	599	4	US-09-343-361-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-073-362-3
: Sequence 3, Application US/09073362
: Patent No. 5942399
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Yue, Henry
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0514 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166

```

TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1665759
; US-09-073-362-3

Alignment Scores:
Pred. No.: 6.69e-113 Length: 515
Score: 1156.50 Matches: 218
Percent Similarity: 69.16% Conservative: 105
Best Local Similarity: 46.68% Mismatches: 141
Query Match: 28.07% Indels: 3
DB: 2 Gaps: 2

US-09-667-170A-440 (1-2239) x US-09-073-362-3 (1-515)
QY 334 GAGAAAGTGGAGTGAAGAGAAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGGC 393
DB 37 GluThrMetGlnLeuLeuLysLysGluIleSerLeuLeuAsnGlyValSerLeuValGly 56
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGGGCTGCTCCAGAACACGGCAGC 453
DB 57 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuValHisThrAlaSer 76
QY 454 GTGGGCATGCTCTGACCACCTTGACGCTGTGTGGGCTGTCTGCTCACTATTGGAGTTTG 513
DB 77 TyrGlyMetSerLeuIleValTrpPalaIleGlyLeuPheSerValValGlyAlaLeu 96
QY 514 TCTTATGCTGAATGGGAACACTATATAAGAAATCTGGAGGTCACTACACATATATTGG 573
DB 97 CysTyrAlaGluLeuGlyThrIleThrLysSerGlyAlaSerTyrAlaTyrIleLeu 116
QY 574 GAAGTCTTTGGTCCATTACCAAGCTTTCTAGAGTCTGGTGGAACTCCCTCATATAATACGC 633
DB 117 GluAlaPheGlyGlyPheIleAlaPheIleArgLeuTrpValSerLeuLeuValGlu 136
QY 634 CTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGAGCTTACATCTGGAACACCTTTT 693
DB 137 ProThrGlyGlnAlaIleAlaIleAlaIlePheAlaAsnTyrIleIleGlnProSerPhe 156
QY 694 ATTCATGTGAATCCCTGAACCTTCGATCAAGCTTACAGCTTACAGCTGTGGGCAATACGTGA 753
DB 157 ProSerCysAspProProTyrLeuAlaCysArgLeuLeuAlaAlaCysIleCysLeu 176
QY 754 GTGATGGTCTTAATAGCATGATGCTGAGCTGGAGCGCCGGATCCAGATTTCTTAAAC 813
DB 177 LeuThrPheValAsnCysAlaTyrValLysTrpGlyThrArgValGlnAspThrPheThr 196
QY 814 TTTTGAAGTCAAGCAATCTTGATAATTATAGTCCCTGGAGTTATGACGATTAATATAA 873
DB 197 TyrAlaLysValValAlaLeuIleAlaIleValMetGlyLeuValLysLeuCysGln 216
QY 874 GGTCAACGACAGAACTTTAAAGACGCTTTTCCAGGAGAGATTCAGTATTACGGGGTTG 933
DB 217 GlyHisSerGlnHisPheGlnAspAlaPheGluGlySerSerTyrAspMetGlyAsnLeu 236
QY 934 CCACTGGCTTTTATTATGAATGATGATGATGCTGGCTGGTTTACCTCAACTTTCTT 993
DB 237 SerLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTrpAspThrLeuAsnPheVal 256
QY 994 ACTCAAGCACTAGAAACCTGAAACCAATCCCTTGCATATATGATATATATCCATGCC 1053
DB 257 ThrGluGluIleLysAsnProGluArgAsnLeuProLeuAlaIleGlyIleSerMetPro 276
QY 1054 ATTGTCACCAATGCTGTGTCACAAATGTCGCTTACTTTACGACCAATTAATGCTGAG 1113
DB 277 IleValThrLeuIleTyrIleLeuThrAsnValAlaTyrTyrThrValLeuAsnIleSer 296

```

```

RESULT 2
US-09-243-920-3
; Sequence 3, Application US/09243920
; Patent No. 5981242
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,920
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/073,362

```

; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cetrone, Michael C
 ; REGISTRATION NUMBER: 39,132
 ; REFERENCE/DOCKET NUMBER: PF-0514 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ;
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 515 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 1665759
 ;
 ; US-09-243-920-3

Alignment Scores:	
Pred. No.:	6,698-113
Score:	136.50
Percent Similarity:	69.18%
Best Local Similarity:	46.68%
Query Match:	26.07%
DB:	2
Length:	515
Matches:	218
Conservative:	105
Mismatches:	141
Indels:	3
Gaps:	2

US-09-667-170A-440 (1-2239) x US-09-243-920-3 (1-515)

334	QY	334	GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGCGAGCTCCATTATCATTCATGGC	393
		Db	: : : : : : : : : : : :	
37	Db	37	GlutThrMetGlnLeuLysLysGluIleSerLeuLeuLeuAsnGlyValSerLeuValValGly	56
394	QY	394	ACCATCATTTGGACAGCAATCTTCATCTCCTTAAGCGCTGCTCAGAACACGGCGAGC	453
		Db	: : : : : : : : : : : :	
5	Db	5	AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuValHisThrAlaSer	76
454	QY	454	GTGGCGCATGCTCTGACCATCTGGACGCTGTGGGCTCTGTCACTATTGTGGAGCTTTG	513
		Db	: : : : : : : : : : : :	
77	Db	77	TyrGlyMetSerLeuIleValTrpAlaIleIleGlyLeuPheSerValValGlyAlaLeu	96
514	QY	514	TCATATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTACACATATATTTTG	573
		Db	: : : : : : : : : : : :	
97	Db	97	CysTyrAlaGluLeuGlyThrThrIleThrLysSerGlyAlaSerTyrAlaTyrIleLeu	116
574	QY	574	GAGCTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGC	633
		Db	: : : : : : : : : : : :	
117	Db	117	GluAlaPheGlyGlyPheIleAlaPheIleArgLeuValSerLeuLeuValValGlu	136
634	QY	634	CTGTCAGCTACTGCTGTCATATCCCTGGCATTTGGACGCTACATCTCGGAACCATTTTTT	693
		Db	: : : : : : : : : : : :	
137	Db	137	ProThrGlyGlnAlaIleIleAlaIleThrPheAlaAsnTyrIleIleGlnProSerPhe	156
694	QY	694	ATTCAATGTGAATCCCTGAAGTTCGGATCAAGCTCATTACAGCTGTGGCGCATCACTGA	753
		Db	: : : : : : : : : : : :	
157	Db	157	ProSerCysAspProProTyrLeuAlaCysArgLeuLeuAlaAlaCysIleCysLeu	176
754	QY	754	GTGATGGTCCTAAATAGCATGAGTGCAGCTCGAGCGCCGGATCCAGATTTTCTTAACC	813
		b	: : : : : : : : : : : :	
177	Db	177	LeuThrPheValAsnCysAlaTyrValLysTrpGlyThrArgValGlnAspThrPheThr	196
814	QY	814	TTTTTCAAGCTCAGCAATCTCTGATATATTAGTCCTCGAGTTATGCAAGCTAATATAA	873
		b	: : : : : : : : : : : :	
197	Db	197	TyrAlaLysValValAlaIleIleAlaIleIleValMetGlyLeuValLysLeuCysGln	216
874	QY	874	GTCAACGCGAGACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAGATTAACCGGTTG	933
		b	: : : : : : : : : : : :	
217	Db	217	GlyHisSerGluHisPheGlnAspAlaPheGluGlySerSerTrpAspMetGlyAsnLeu	236
934	QY	934	CCACTGCTTTTTTATTATGGATGATATGCATATGCTGGCTGGTGTTTTAOCTCAACTTTGTT	993
		b	: : : : : : : : : : : :	
237	Db	237	SerLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTrpAspThrLeuAsnPheVal	256

```

RESULT 3
US-09-073-362-1
: Sequence 1, Application US/09073362
: Patent No. 5942399
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Yue, Henry
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: incyte pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible

```

```

: RESULT 4
: US-09-243-920-1
: ; Sequence 1, Application US/09243920
: ; Patent No. 5981242
: ; GENERAL INFORMATION:
: ; APPLICANT: Hillman, Jennifer L.
: ; APPLICANT: Yue, Henry
: ; APPLICANT: Cortey, Neil C.
: ; TITLE OF INVENTION: AMINO ACID
: ; NUMBER OF SEQUENCES: 3
: ; CORRESPONDENCE ADDRESS:
: ; ADDRESSEE: Incyte Pharmaceuticals

```


STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,920
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/073,362
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Certione, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PP-0514 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 511 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: ESOGTUT02
CLONE: 2667831
US-09-243-920-1

Alignment Scores:
Pred. No.: 2,85e-110 Length: 511
Score: 1131.50 Matches: 214
Percent Similarity: 68.00% Conservative: 109
Best Local Similarity: 45.05% Mismatches: 149
Query Match: 27.46% Indels: 3
DB: 2 Gaps: 2

US-09-667-170A-440 (1-2239) x US-09-243-920-1 (1-511)

QY 310 GGAAGACGCTTTCAGGAGAGAGAAAGTGCAGCTCAAGAGAAAGTCACTTTACTG 369
Db 21 GlyAspGlyAlaSerProGlyProGluGlnValLysLeuLysLysGluLeuSerLeu 40
QY 370 AGGGAGTCTCCATTATTCATTCAGGACCAATTCATTCAGGAGGAGAAATCTCTCTCAAG 429
Db 41 AsnGlyValCysLeuLeuValGlyAsnMetIleGlySerGlyIlePheValSerProLys 60
QY 430 GCGGTGCTCCAGACAGGCGAGCGTGGGCAATGCTCTGACCATCTGGACGCTGGTGGG 489
Db 61 GlyValLeuIleTyrSerAlaSerPheGlyLeuSerLeuValIleTyrAlaValGlyGly 80
QY 490 GTCCCTGTCACTATTTGGAGCTTTGCTATCTGCTGAATGGAGCACTATAAAGAAATCT 549
Db 81 LeuPheSerValPheGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleLysSer 100
QY 550 GGAGGTCACTACACATATTTTGAAGTCTTTGGTCCATTACCACTTTGTACCAAGTC 609
Db 101 GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu 120
QY 610 TGGGTGAACCTCCTCAATACGCTCCGAGCTACTGCTGTATATCCCTGGCAATTTGA 669
Db 121 TrpThrSerLeuIleGluProThrSerGlnAlaIleAlaIleThrPheAla 140
QY 670 CCTACATCTCGAACCACTTTTATTCAATGTGAATCCCTGAACCTTCGGATCAAGCTC 729
Db 141 AsnTyrMetValGlnProLeuPheProSerCysPheAlaProTyrAlaAlaSerArgLeu 160

QY 730 ATTACAGCTGTGGCATAAAGTCTAGTGATGCTCTAATAAGCATGAGTGTACGTGGAGC 789
Db 161 LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyrValLysTrpGly 180
QY 790 GCCCGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATCTGATAATTTATATGTC 849
Db 181 ThrLeuValGlnAspIlePheThrTyrAlaLysValLeuAlaLeuIleAlaValIleVal 200
QY 850 CTTGAGTTATGTCAGCTAAATTAAGGTCAACGACAGCACTTTAAAGACCCCTTCACAGA 909
Db 201 AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGluGly 220
QY 910 AGAGATTCAAGTATTACCGGTTGGCCACTGGCTTTTATTATGGAAATGTATCATATGCT 969
Db 221 SerSerPheAlaValGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer 240
QY 970 GGCTGGTTTACCTCAACTTTTACTGAAGAGTAGAAACCCCTGAAACCAACCAATCC 1029
Db 241 GlyTrpAspThrLeuAsnTyrValThrGluGluIleLysAsnProGluArgAsnLeuPro 260
QY 1030 CTTGCAATATGATATCCATGCGCATTCACCATTTGGCTATGCTGACAAATGTGGCC 1089
Db 261 LeuSerIleGlyIleSerMetProIleValThrIleIleTyrIleLeuThrAsnValAla 280
QY 1090 TACTTTACGACCACTAATCTGAGGAGCTGCTGCTTTCAAATGCAGTGGCAGTGCCTTT 1149
Db 281 TyrTyrThrValLeuAspMetArgAspIleLeuAlaSerAspAlaValAlaValThrPhe 300
QY 1150 TCTGAGCGCTACTGGAAATTTCTCATTAGCAGTTCCGATCTTTGGCTCTCTCTCGA 1209
Db 301 AlaAspGlnIlePheGlyIlePheAsnTrpIleIleProLeuSerValAlaLeuSerCys 320
QY 1210 TTTGGCTCCATCAACGCTGCTGCTCTCTCTCTCCAGTTATTTCTATGTTGGCTCGA 1269
Db 321 PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuPhePheValGlySerArg 340
QY 1270 GAGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCGCAGACACACTCTCTACCA 1329
Db 341 GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro 360
QY 1330 GCTGTTATTGTTTCCACCCCTTTGACAAATGATGCTCTCTCTGAGACCTCGACAGT 1389
Db 361 SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln 380
QY 1390 CTTTGAATTTCCAGTGTTCAGGCTGCTTTTATTTGGCTGGCAGTCTGCTGGCTG 1449
Db 381 LeuIleAsnTyrTyrSerPheSerTyrTrpPhePheValGlyLeuSerIleValGlyGln 400
QY 1450 ATTATCTTCGATACAAATGCCAGATATGCATGCTCTTCAAGTGGCAGCTGTTTCATC 1509
Db 401 LeuTyrLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerValPhePhe 420
QY 1510 CCAGCTTTGTTTCTCTTCACATGCTCTTCATGTTGGCTTTCCTCTTATTCGGAACCA 1569
Db 421 ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr 440
QY 1570 TTTAGTACAGGATGGCTTTCGTCATCTGACTGGAGTCCCTGCTGCTATTTATCTCTT 1629
Db 441 IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIle 460
QY 1630 ATT-----ATATGGCACAAGAAACCCAGGTGTTTACAAATAATGTACAGAAATAACC 1683
Db 461 IleArgValProGluHisLysArgProLeuTyrLeuArgIleValGlySerAlaThr 480
QY 1684 AGAACATTTCAATAATA---CTGGAAGCTTTGACCCAGAGAAGAT 1725
Db 481 ArgTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetAsp 495

RESULT 5

US-08-825-781-3

; Sequence 3, Application US/08825781

; Patent No. 5843727

us-09-667-170a-440.n2p.rai

Thu Apr 17 09:48:24 2003

GENERAL INFORMATION:
 APPLICANT: Hillman, Jennifer L.
 TITLE OF INVENTION: HUMAN TUMOR-ASSOCIATED MEMBRANE PROTEIN
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/825,781
 FILING DATE: Herewith
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0262 US
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:

INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 241 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 392990
 US-08-825-781-3

Alignment Scores:
 Pred. No.: 1e-56
 Score: 621.00
 Percent Similarity: 68.49%
 Best Local Similarity: 47.90%
 Query Match: 15.07%
 Gaps: 0
 Db: 2

US-09-667-170A-440 (1-2239) x US-08-825-781-3 (1-241)

QY 1009 AACCTGAAACCAATCCCTGCAATATGATATCATGCCATTCACCATGGC 1068
 Db 3 AsnProTyrArgAsnLeuProLeuAlaIleSerLeuProIleValThrLeuVal 22
 QY 1069 TATGCTGCAATATGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTTCCTCA 1128
 Db 23 TyrValLeuThrAsnLeuAlaTyrPheThrLeuSerThrAsnGlnMetLeuThrSer 42
 QY 1129 AATGAGTGGCAGTGACCTTTCTGACGGCTACTGGGAATTTCTCATAGCAGTTCCG 1188
 Db 43 GluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleLeuPro 62
 QY 1189 ATCTTTGTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
 Db 63 ValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThrSerArg 82
 QY 1249 TTATTCTATGTCGCTCGAGAGGTCACCTTCACGAATTCCTCTCCATGATTCATGTC 1308
 Db 83 LeuPhePheValGlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisPro 102

QY 1309 CGCAAGCACATCTCTACACGCTGTTATTGTTGACACCTTTGACAAATGATAATGCTC 1368
 Db 103 GlnLeuLeuThrProValProSerLeuValPheThrCysValMetThrLeuMetTyrAla 122
 QY 1369 TTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTTCAGTTTTCAGGTGGCTTTTATT 1428
 Db 123 PheSerArgAspIlePheSerIleAsnPhePheSerPhePheAsnTrpLeuCysVal 142
 QY 1429 GGCTGCGAGTTCGCTGCTGATTTCTCGATACAAATGCCAGATATCATCGCTCCT 1488
 Db 143 AlaLeuAlaIleLeuGlyMetMetTrpLeuArgPheGlyLysProGluLeuGluArgPro 162
 QY 1489 TTCAAGTGCCACTGTTTCATCCAGCTTTGTTTCTTCCATCATGCTCTTTCATGCTTCC 1548
 Db 163 IleLysValAsnLeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAla 182
 QY 1549 CTTTCCTCTATTCGACCACTTTAGTACAGGGATTGGCTTCGTCATCATCTGACTGGA 1608
 Db 183 ValSerPheTrpLysThrProLeuGluCysGlyIleGlyPheAlaIleIleLeuSerGly 202
 QY 1609 GTCCCTGCTATTATCTCTTTTATTATATATGGCAAGAACCCAGGTGGTTTGAATAATG 1668
 Db 203 LeuProValTyrPhePheGlyValTrpTrpLysAsnLysProLysTrpIleLeuGlnVal 222
 QY 1669 TCAGAGAAATACCAAGAACATTTACAAATATATGCTGAGTTCGTACCAAGAA 1722
 Db 223 IlePheSerValThrValLeuCysGlnLysLeuMetGlnValProGlnGlu 240

RESULT 6
 US-08-825-781-4
 ; Sequence 4, Application US/08825781
 ; Patent No. 5843727
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; TITLE OF INVENTION: HUMAN TUMOR-ASSOCIATED MEMBRANE PROTEIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/825,781
 FILING DATE: Herewith
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0262 US
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 241 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 181908

US-08-825-781-4

Alignment Scores:

Pred. NO.:	4, 29a-56	Length:	241
Score:	615.00	Matches:	115
Percent Similarity:	68.07%	Conservative:	47
Best Local Similarity:	48.3a	Mismatches:	76
Query Match:	14.93%	Indels:	0
DB:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x US-08-825-781-4 (1-241)

1009	Y	1009	AA	CCCTGAAAAAACCAATTC	CCCTTGC	CAATATGTATATCCATGCCATTTGTACCAATGGC	1068
					:::		
3	db	3	Asn	ProTyrArgAsnLeu	ProLeuAlaIleIle	SerLeuProIleValThrLeuVal	22
1069	Y	1069	TAT	TGCTGCACAAATGTGGCT	ACTTTT	TACGACCAATTAATGCTGAGGAGCTGCTGCTTTCA	1128
	db		Tyr	ValLeuThrAsnLeuAlaTyr	PheThrThrLeuSerThrGluGlnMetLeuSerSer		42
1129	Y	1129	AAT	GCGAGTGGCAGTCACCT	TTTCTG	AGCGCTACTGSGGAATTTCTCATAGCAGTTCG	1188
43	db	43	Glu	AlaValAlaValAsp	PheGlyAsnTyrHisLeuGlyValMetSerTrpIleLeuPro		62
1189	Y	1189	ATC	TTTGTGGCCCTCTCT	CTGTTT	GCGTCATGAACGGTGGTGTGTGCTGCTCCAGC	1248
63	db	63	Val	PheValGlyLeuSerCysPhe	GlySerValAsnGlySerLeuPheThrSerSerArg		82
1249	Y	1249	TAT	TCTATGTGGTGCTC	GAGAGGCTCACCTT	CCAGAAATCCTCCATGATTCATGTC	1308
83	db	83	Leu	PhePheValGlySerArg	GluGlyHisLeuProSerIleLeuSerMetIleHisPro		102
1309	Y	1309	CGC	AAGCACCTCTCT	TACCAGCTGTTATGTT	TGTCACCTTTGGACAATGATATGCTC	1368
103	db	103	Gln	LeuLeuThrProValPro	SerLeuValPheThrCysValMetThrLeuLeuTyrAla		122
1369	Y	1369	TTCTCTG	GAGACTCGACAGCT	TTTTTG	AATTTCTCTCAGTTTCCAGGTCGCTTTTAT	1428
123	b	123	Phe	SerTyrAspIlePheSer	ValIleAsnPhePheSerPheAsnTrpLeuCysVal		142
1429	Y	1429	GGGCTGG	CAGTGTGGCGTGATTT	ATCTTCC	ATACAAATGCCAGATATGATCGTCT	1488
143	b	143	Ala	LeuAlaIleIleGlyMet	IleTrpLeuArgHisArgLysProGluLeuGluArgPro		162
1489	Y	1489	TTCAAGTGG	CCATGTTTCATCC	CCAGCTTTGTTT	TCCCTTCATGCGCTCTCTCATGGTCC	1548
163	db	163	Ile	LysValAsnLeuAlaLeu	ProValPhePheIleLeuAlaCysLeuPheLeuIleAla		182
1549	Y	1549	CTTTCC	CTCTATTTCGACCACT	TATTAGTACAGGATTTGGCTTCGTCAT	CACTCTCACTGGA	1608
183	db	183	Val	SerPheTrpLysThrPro	ValGluCysGlyIleGlyPheThrIleIleLeuSerGly		202
1609	Y	1609	GTCCCTCG	TATATCTCTTTAT	TATATGGGCAAGAAACCCAGCTGGTGTAG	AATATG	1668
203	db	203	Leu	ProValTyrPhePheGly	ValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGly		222
1669	Y	1669	TCAGAGAA	AATAACCAACATTA	CAATAATACTCGAAGTTGTACCAAGAA		1722
223	db	223	Ile	PheSerThrThrValLeu	CysGlnLysLeuMetGlnValValProGlnGlu		240

RESULT 7

US-08-825-781-1
; Sequence 1, Application US/08825781
; Patent No. 5843727

: GENERAL INFORMATION:
 : APPLICANT: Hillman, Jennifer L.
 : APPLICANT: Shah, Purvi
 : TITLE OF INVENTION: HUMAN TUMOR-ASSOCIATED MEMBRANE PROTEIN
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 3174 Porter Drive
 : CITY: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/825,781
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36, 749
 REFERENCE/DOCKET NUMBER: PF-0262 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:

FORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 245 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: PANTUT02
 CLONE: 2236771
 825-781-1

Alignment Scores:

Pred. No.:	1.44e-46	Length:	245
Score:	524.50	Matches:	102
Percent similarity:	67.25%	Conservative:	52
Best Local Similarity:	44.54%	Mismatches:	72
Query Match:	12.73%	Indels:	3
DB:	2	Gaps:	2

US-09-667-170A-440 (1-2239) x US-08-825-781-1 (1-245)

	OY	1048	ATGCCCATTTGACCAATTGGCTATGTGCTGACAATAAGTGGCTACTTTACGACCATTAAAT	1107
	Dd	1	MetProIleValThrIleIleIleIleLeuThrasnValalaTyrrTrhrValLeuAasp	20
	OY	1108	GCTGAGGAGTGCTGCTTCCTCAAATGGCAGTGCGCAGTGCACCTTTTTCTGACGGCTACTGGGA	1167
	Dd	21	MetArgAspIleLeualaSerasnAlavalalavalThrPheAlaaSpcInilePheGly	40
	OY	1168	AATTTCTCATTAAGCCATTCGATCTTTGTGGCCTCTCTGCTTTGGCTTCATCAACCGGT	1227
	Dd	41	IlePheAsnTrpIleIleProLeuSerValAlalaLeuSerCysPheGlyGlyLeuAsnAla	60
	OY	1228	GGTGTCTTTGGCTCTCCAGGTATTCTATGTTCGTCTCGAGAGGTGCACCTTCAGAA	1287
	Dd	61	SerIlevalAlalaSerArgLeuPhePheValGlySerArgGluGlyHisLeuProasp	80
	OY	1288	ATCCTCTCATGATTCATGTCGCCAACACACTCTCTACCAAGCTGTTATGTTTGGCAC	1347
	bD	81	AlaIleCysMetIleHisValGlutArgPheThrProValProSerLeuLeuPheAsnGly	100
	OY	1348	CCTTTGACAATGAATGAAGCTCTTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCCCTCAGT	1407
	bD	101	IleMetAlaLeuIleLeyrLeuCysValGluAspIlePheGlnLeuIleasntyryrSer	120
	OY	1408	TITGGCAGGTGGCTTTTATTGGGGCTGGCAGTGTGGGGCTGATTTATCTTCGCATACAAA	1467
	b	121	PheSertyrTrpPhePheValGlyLeuSerIleValGlyGlnIleutyryleuargtrplrys	140
	y	1468	TGCCCATATGCATCGCTCTTTCAGGTGGCAGTGTTCATCCAGCATTGTGTTCCTTC	1527

us-09-667-170a-440.n2p.ra1

Thu Apr 17 09:48:24 2003

Db 141 GluProAspArgProArgProLeuTysLeuSerValPhePheProLeuValPheCysLeu 160
 Qy 1528 ACATGCTCTTCATGTTGCTCCCTTCCTCTATTCGAGCCCATTTAGTACAGGGATTCGC 1587
 Db 161 CysThrIlePheLeuValAlaValProLeuTyrSerAspThrIleAsnSerLeuIleGly 180
 Qy 1588 TTCGTCATCATCTGCTGAGTCCCTGCTGCTATTCCTCTTATT-----ATATGGGAC 1641
 Db 181 IleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIleIleArgValProGluHis 200
 Qy 1642 AAGAAACCCAGGTGTTTACGAAATATCTCAGAGAAATAACACAGAACATTTACAAATAATA 1701
 Db 201 LysArgProLeuTyrLeuArgGlyIleValGlySerAlaThrArgTyrLeuGlnValLeu 220
 Qy 1702 ---CTGGAAGTGTACCAAGAGAAGAT 1725
 Db 221 CysMetSerValAlaAlaGluMetAsp 229

RESULT 8
 US-08-132-990A-8
 ; Sequence 8, Application US/08132990A
 ; Patent No. 5834589
 ; GENERAL INFORMATION:
 ; APPLICANT: MERUELO, DANIEL
 ; APPLICANT: YOSHIMOTO, TAKAYUKI
 ; TITLE OF INVENTION: Human Retrovirus Receptor and DNA Coding Therefor
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.24
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/132,990A
 ; FILING DATE: 07-OCT-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/084,729
 ; FILING DATE: 29-JUN-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/05569
 ; FILING DATE: 11-JUN-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/899,075
 ; FILING DATE: 11-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/806,178
 ; FILING DATE: 13-DEC-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/627,950
 ; FILING DATE: 14-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mistrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 8105-004-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9864
 ; TELEX: 66441 PENNIE
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 629 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-132-990A-8

Alignment Scores: 2.93e-21 Length: 629
 Pred. No.: 286.00 Matches: 103
 Score: 43.16% Conservativity: 80
 Percent Similarity: 24.29% Mismatches: 183
 Best Local Similarity: 6.94% Indels: 58
 Query Match: 2 Gaps: 12
 DB:
 US-09-667-170A-440 (1-2239) x US-08-132-990A-8 (1-629)
 Qy 331 AGAGAGAAAGTCGAGCTGAAGAGAAAGTCATCTTACTAGGGAGTCTCCATTTATCAT 390
 Db 23 ArgGluThrArgLeuSerArgCysLeuAsnThrPheAspLeuValAlaLeuGlyVal 42
 Qy 391 GGCACCATCATTTGAGCAGCAATCTTCATCTCTCTAAGGGCTGCTCCAGAACACGGGC 450
 Db 43 GlySerThrLeuGlyAlaGlyValTyrValLeuAlaGlyAlaValAlaArgGlu 60
 Qy 451 AGCGTGGCATGTCTCTGACCATC---TGGACGGTGTGGGTCTCTGCTCACTATTTCGA 507
 Db 61 AsnAlaGlyProAlaIleValIleSerPheLeuIleAlaAlaLeuAlaSerValLeuAla 80
 Qy 508 GCTTTGCTTATGCTGAATTTGGGAACAATATAAGAAATCTGGAGTCTTACACATAT 567
 Db 81 GlyLeuCysTyrGlyGluPheGlyAlaArgValProLysThrGlySerAlaTyrLeuTyr 100
 Qy 568 ATTTTGGAGTCTTTGGTCCATTACACGCTTTTTCAGACTCTGGTGGGAGTCTCTC--- 624
 Db 101 SerTyrValThrValGlyGluLeuTyrAlaPheIleThrGlyTyrAsnLeuIleLeuSer 120
 Qy 625 ---ATAATACGCCCTGACGCTACTCTCT---GTGATATCCCTGGCATTTGGACGCTACAT 678
 Db 121 TyrIleIleGlyThrSerSerValAlaAlaArgAlaThrPheAspGluLeuIle 140
 Qy 679 CTGGAACCATTT-----TTTATTCAA 699
 Db 141 GlyArgProIleGlyGluPheSerArgThrHisMetThrLeuAsnAlaProGlyValLeu 160
 Qy 700 TGTGAATCCCTGAACTTCGATCAGCTCATTACAGCTGTGGCATCACTGTAGTATGATG 759
 Db 161 AlaGluAsnProAspIlePheAlaValIleIle-----IleLeuIleLeuThr 176
 Qy 760 GTCTTAATAGCATGATGCTCAGCTGGAGCCCGCATCCAGATTTTCTTAACCTTTTGC 819
 Db 177 GlyLeuLeuThrLeuGlyValLysGluSerAlaMetValAsnLysIlePheThrCysIle 196
 Qy 820 AAGCTCACAGCAATCTGATATATATTAGTCCCTGGAGTTATCCAGCTAATAAAGGTCAA 879
 Db 197 AsnValLeuValLeuGlyPheIleMetValSerGly-----PheValLysGlySer 213
 Qy 880 ACGCAGAACTTTAAA----- 894
 Db 214 ValLysAsnThrGlnLeuThrGluGluAspPheGlyAsnThrSerGlyArgLeuCysLeu 233
 Qy 895 -----CACGCCCTTTTCAGGAAGAGATTCAAGTATTACGGGTGGCCACTGGCTTTTAT 948
 Db 234 AsnAsnAspThrLysGluGlyLysProGlyValGlyGlyPheMetProPheGlyPheSer 253
 Qy 949 TATGGAATG-----TATGCATATGCTGGCTGGTGGTTTACCTCAAC 987
 Db 254 GlyValLeuSerGlyAlaAlaThrCysPheTyrAlaPheValGlyPheAspCysIleAla 273
 Qy 988 TTTGTTACTGAAGAGTAGAAGAACCCCTGAAAAACCATTCCTCCCTTGCATATATATATCC 1047
 Db 274 ThrThrGlyGluGluValLysAsnProGlnLysAlaIleProValGlyIleValAlaSer 293
 Qy 1048 ATGGCCATTGTCCACCATTTGGCTATGCTGACAAATCTGGCCTACTTTACGACCATTAAT 1107
 Db 294 LeuLeuIleCysPheIleAlaTyrPhe-----GlyValSerAlaAlaLeuThrLeuMet 311
 Qy 1108 GCTGAGGAGCTGCTCTCTTTCAAATGTCAGTGGCAGTGGACCTTTCTGACGGCTACTGGA 1167

QY 1276 CACCTTCCAGAAATCTCTCCATGATTCATGTCGCGGACGACACCTCTACAGCTGT 1335
 Db 370 LeuLeuPheLysPheLeuAlaAsnValAsnArgThrLysThrProfilLeAlaThr 389
 QY 1336 ATTGTTTGGACCCCTTGCACAAATGATGCTCTCTCTGGAGACCTCGACAGCTTTTG 1395
 Db 390 LeuAlaSerGlyAlaValAlaAlaValMetAlaPheLeuPheAspLeuLysVal 409
 QY 1396 AATTTCTCAGTTTGGCAGGTGCTTTTATTTGGCTGGCAGTCTCTGGCTGATTTAT 1455
 Db 410 AspLeuMetSerIleGlyThrLeuLeuAlaTyrSerLeuValAlaAlaCysValLeuVal 429
 QY 1456 CTTGATACAAA 1467
 Db 430 LeuArgTyrGln 433
 RESULT 10
 US-09-134-001C-4290
 ; Sequence 4290, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134.001C
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 4290
 ; LENGTH: 521
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-4290
 Alignment Scores:
 Pred. No.: 6.2e-21 Length: 521
 Score: 282.50 Matches: 135
 Percent Similarity: 43.39% Conservative: 108
 Best Local Similarity: 24.11% Mismatches: 206
 Query Match: 6.86% Indels: 111
 DB: 4 Gaps: 26
 US-09-667-170A-440 (1-2239) x US-09-134-001C-4290 (1-521)
 QY 127 TTTTAT---CATATCTGGATTTTGGATTCCTTTGTTTCTCATCATCTGATTCAGGA 183
 Db 8 PheTyrLeuHisValAspPheLeuLeuAlaArgIleValLysGluCysLeuLeuThr--- 26
 QY 184 AAGCTGTGTGTCACCATCTCCAAAGAGAGGTTTACCTGAGGGAATGTTACGGGAGG 243
 Db 27 ---HisValAsnSerLysLeuMetLeuTyrArgPheValMetGlySerPhePheAsnArg 45
 QY 244 CTGCTCTCCCTGGGCAACAGGACCCACCTGGGCGAGGACGCTTTTCAGGAAGAGAGC 303
 Db 46 MetThr-----ArgLysGluAsnPro-----ThrIle 54
 QY 304 TTTTTCAGGAAGAGAGCGCTTTTTCAGGAAGAGAGAGAGTGCAGTGAAGAGAAAGTCACT 363
 Db 55 TyrGlnSerLysAspGly-----HisLeuLysArgThrLeuArg 67
 QY 364 TTACTGAGGGAGTCTCATATATCATTCGACCATCATTTGGAGCAGGAATCTTCATCTCT 423
 Db 68 ValArgAspPheLeuAlaLeuGlyValGlyThrIleValSerThrSerIlePheThrLeu 87
 QY 424 CTTAAGGGCTGCTCCAGCAACACGGGCGAGCGTGGGCGATGCTCTGACCATCTGACCGGTG 483
 Db 88 ProGlyValValAlaAlaGluHisAlaGlyProAlaValSerLeuSer---PheLeuLeu 106

QY 484 TGTGGGCTCTCTCCTACTATTGAGCTTTGCTTATGCTGAATTTGGGAACAACATATAAG 543
 Db 107 AlaAlaIleValAlaGlyLeuValAlaPheThrTyrAlaGluMetAlaSerThrMetPro 126
 QY 544 AATCTGGAGGTCAATACACATATATTTTGGAGTCTTTGGTCCATTTACAGCTTTTCTA 603
 Db 127 PheAlaGlySerAlaTyrSerTrpIleAsnValLeuPheGlyGluLeuPheGlyTrpVal 146
 QY 604 CGAGTCTGGTGAACCTCTCATATACCCCTGCGAGCTACTGCTGTGATATCCCTGCGCA 663
 Db 147 AlaGlyTrpAla---LeuLeuAlaGluTyrPheIleAlaValAlaPheValAlaSerGly 165
 QY 664 TTTTGA-----CGCTACATTTCTGGAACCATTTTATTCAATGTGAAATCCCTGAA 714
 Db 166 PheSerAlaAsnLeuArgGlyLeuIleAlaProLeuGlyIle-----SerLeuProLys 183
 QY 715 CTT-----GCCATCAAGCTCATACAGCTGTGGGC 744
 Db 184 SerLeuSerAsnProPheGlySerAsnGlyGlyValIleAspIleIleAlaValVal 203
 QY 745 ATAATCTAGTATGTCCTAAATAGCATGAGTGTGAGTGGAGCCCGGATCCAGATT 804
 Db 204 IleIleLeuThrAlaLeuLeuSerArgGlyMetAsnGluAlaAlaArgMetGluAsn 223
 QY 805 TTCTTAACCTTTTGCAGCTCACAGCAATTTCTGATATATTAGTCTCTGAGTATGAG 864
 Db 224 ValLeuValIleLeuLysValLeuAlaIleIleLeuPheValIleValGlyLeuThrAla 243
 QY 865 CTAATTAAGGTCAACCCAGCACTTTAAAGAC----- 897
 Db 244 Ile-----AsnSerAsnTyrIleProPheIleProGluHisLys 257
 QY 898 -----GCCTTTTCAGGAAGATTCACAGTATTACGCGTGGCTGGCT 942
 Db 258 ValThrGluThrGlyAspPheGlyGly----- 266
 QY 943 TTTTATTGGAATGATTCATATGCTGCTGCTGTTTACCTCAACTTTGTT----- 993
 Db 267 ---TrpGlnGlyIleTyrAlaGlyValSerMetIlePheLeuAlaTyrIleGlyPheAsp 285
 QY 994 -----ACTGAAGAAGTAGAAACCCCTGAAAAACCATCCCTTGGCAATA 1038
 Db 286 SerIleAlaAlaAsnSerAlaGluAlaIleAsnProGlnLysThrMetProArgGlyIle 305
 QY 1039 TGTATATCATGCGCCATT-----GTCACTATGCTGATGCTGCTGACAAAT 1083
 Db 306 LeuGlySerLeuIleValAlaIleValLeuPheValAlaValAlaLeuValLeuValGly 325
 QY 1084 GTGCGCTACTTTACGACCATTAATGCTGAGAGCTGCTCTTCAATGCGAGTGGCAGTG 1143
 Db 326 MetPheHisTyrSerGln-----TyrAlaAspAsnAlaGluProVal 339
 QY 1144 ACCTTTTCTGAGGGCTACTGGGAATTTCTCATTAGCA-----GTTCCGATCTTTGTT 1197
 Db 340 GlyTrpAlaLeuArgGlyHisGlyIleIleAlaIleValGlnAlaIleSer 359
 QY 1198 GCCCTCTCTCTGCTGCTGCTCATGAGCGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1257
 Db 360 ValIleGlyMetPheThrAlaLeuIleGlyMetLeuAlaGlySerArgLeuLeuTyr 379
 QY 1258 GTTGGCTCTGAGAGGTGCTACCTCCAGAAATCTCTCCATGATTCATGCTGCGCAAGCAC 1317
 Db 380 SerPheGlyArgAspGlyLeuLeuProSerTrpLeuSerGlnLeuAsn---HisLysHis 398
 QY 1318 ACTCTCTTACCATGCTGTTATTTGTTTTCACCCCTTTGACAAATGATGCTCTCTCTGGA 1377
 Db 399 LeuProAsnArgAlaLeuAlaIle-----LeuThrIleIleGlyValIleGly 415
 QY 1378 GAC-----CTCGACAGTCTTTTGAATTTCTCACTTTGCGAGGTGG 1419
 Db 416 SerMetPheProPheAlaPheLeuAlaGlnLeuIleSerAlaGlyThrLeuValAlaPhe 435
 QY 1420 CTTTTTATTGGCTGGCAGTGTGCTGGCTGATTTATCTTCGATACAAATGCCCGATATG 1479

Db 436 MetPheValSerLeuAlaMetTyrArgLeu-----ArgLysArgGluGlyLysAspLeu 453
 QY 1480 CATCGTCTCT---TTCAAGGTGCCACTGTTCT---ATCCAGCTTTGTTTCTTCTCACA 1530
 Db 454 ProLysProGluPheLysLeuProLeuTyrProLleLeuProAlaLeile---ThrPheLe 472
 QY 1531 TGCCTCTCATGTTGCTCCCTTCCCTC---TATTCGGAGCCCATTTAGTACAGGGATTGCG 1587
 Db 473 LeuValLeuLeuValPheTrpGlyLeuSerPheGluAlaLysLeuTyrThrLeuLeiletrp 492
 QY 1588 TTGCTCATCTCTGACTGAGTCCCTGCTGCTATTTCTCTTTATTTATATATGAGGACAGAA 1647
 Db 493 PheLeuVal-----GlyLeileLeileTyrLeuLeileTyrGlyLeArgHisSerLys 509

RESULT 11

US-08-132-990A-4
 ; Sequence 4, Application US/08132990A
 ; Patent No. 5834589
 ; GENERAL INFORMATION:
 ; APPLICANT: MERUELO, DANIEL
 ; TITLE OF INVENTION: YOSHIMOTO, TAKAYUKI
 ; NUMBER OF SEQUENCES: Human Retrovirus Receptor and DNA Coding Therefor
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.24
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/132,990A
 ; FILING DATE: 07-OCT-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/084,729
 ; FILING DATE: 29-JUN-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/05569
 ; FILING DATE: 11-JUN-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/899,075
 ; FILING DATE: 11-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/806,178
 ; FILING DATE: 13-DEC-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/627,950
 ; FILING DATE: 14-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 8105-004-999
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864
 ; TELEX: 66441 PENNIE
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 622 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-132-990A-4

Alignment Scores:

Pred. No.: 1.41e-20 Length: 622

Score: 279.50 Matches: 102
 Percent Similarity: 42.82% Conservative: 80
 Best Local Similarity: 24.00% Mismatches: 176
 Query Match: 6.78% Indels: 67
 DB: 2 Gaps: 13
 US-09-667-170A-440 (1-2239) x US-08-132-990A-4 (1-622)
 QY 331 AGAGAGAAAGTCAGCTGAAGAGAAAGTCACCTTTACTGAGGGAGTCTCCATTATCAT 390
 Db 23 ArgGluGluSerArgLeuSerArgCysLeuAsnThrTyrAspLeuValAlaLeuGlyVal 42
 QY 391 GGACCATCATTTGAGCAGCAATCTTCATCTCTCTCTAGGGCGTCTCTCAGAACACGGC 450
 Db 43 GlySerThrLeuGlyAlaGlyValTyrValLeuAlaGlyAlaValAlaArgGlu----- 60
 QY 451 AGCGTGGGATGCTCTGTGACCATC---TGGACGGTGTGGGGTCTGTCTCACTATTTTGA 507
 Db 61 AsnAlaGlyProAlaLeileValIleSerPheLeuLeileAlaAlaLeuAlaSerValLeuAla 80
 QY 508 GCTTTGCTCTTATGCTGAATTTGGGAACAATATAAGAAATCTGGAGTCTATACACATAT 567
 Db 81 GlyLeuCysTyrGlyGluPheGlyAlaArgValProLysThrGlySerAlaTyrLeuTyr 100
 QY 568 ATTTTGAAGTCTTTGCTGTCATTTACACAGTCTTTGTACAGTCTGTGGTGAACCTCTC--- 624
 Db 101 SerTyrValThrValGlyGluLeuTrpAlaPheleThrGlyTrpAsnLeuLeuLeuSer 120
 QY 625 ---ATAATACGCGCTGACCTACTGCT---GTGATATCCCTGGGCAATTTGGACGCTACAT 678
 Db 121 TyrIleleGlyThrSerSerValAlaArgAlaTrpSerAlaThrPheAspGluLeuLe 140
 QY 679 CTGGAACCATTT-----TTTATTCAA 699
 Db 141 GlyLysProIleGlyGluPheSerArgGlnHisMetAlaLeuAsnAlaProGlyValLeu 160
 QY 700 TGTGAATCCCTGAACTTGGCTGCACTCACTGCTGAGCTGCTGGGCACTGAGTGTGATG 759
 Db 161 AlaGlnThrProAspIlePheAlaValIlele-----IleleleleLeuThr 176
 QY 760 GTCCTAAATAGCATGAGTGTGAGCTGGAGCGCGGATCCAGATTTTCTTAACCTTTTGC 819
 Db 177 GlyLeuLeuThrLeuGlyValLysGluSerAlaMetValAsnLyslePheThrCysIle 196
 QY 820 AAGCTCACAGCAATTCGATAATATAGTCCCTGGAGTATTCAGCAGCTAAATAAGGTCAA 879
 Db 197 AsnValLeuValLeuLysPheIleValValSerGly-----PheValLysGlySer 213
 QY 880 ACGCAGAACTTTAAAGACGCTTTTTCAGGAGAGATTCAGTATTTACGCGGTTGCCACTG 939
 Db 214 IleLysAsnTrpGln-----LeuThrGluLysAsnPheSerCysAsnAsnAspThr 231
 QY 940 GCTTTTATTATGGA----- 954
 Db 232 AsnValLysTyrGlyGluGlyGlyPheMetProPheGlyPheSerGlyValLeuSerGly 251
 QY 955 -----ATGTATGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002
 Db 252 AlaAlaThrCysPheTyrAlaPheValGlyPheAspCysIleAlaThrThrGlyGluGlu 271
 QY 1003 GTAGAAACCTGAAACCAACCATTTCCCTTGCATATATGATATTCATCCATGCTGCTCACC 1062
 Db 272 ValLysAsnProGlnLysAlaIleProValGlyIleValAlaSerLeuLeuIleCysPhe 291
 QY 1063 ATTGGCTATGCTGCTGACAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122
 Db 292 Ile-----AlaTyrPhe-----GlyValSerAlaAlaLeuThrLeu 303
 QY 1123 CTTTCAATGTCAGTGGAGTACCTTTTCTGAGCGGCTACTGGGAATTTTC----- 1173
 Db 304 MetMetProTyrPheCysLeuAspIleAspSerProLeuProGlyAlaPheLysHisGln 323
 QY 1174 -----TCATTAGCAGTTCGATCTTTTGTGTCCTCTCTCTGCTTT 1212

Db 362 GlyLeuLeuPheLysPheLeuAlaLysIleAsnAsnArgThrLysThrProValIleAla 381
 QY 1333 GTATTGTTTGGACCCCTTGGACAATGAATAGCTCTCTCTGGAGACCTCGACAGCTTT 1392
 Db 382 ThrValThrSerGlyAlaIleAlaValMetAlaPheLeuPheGluLeuLysAspLeu 401
 QY 1393 TTGAATTTCTCAGTTTGGCCAGTGGCTTTTATTGGCTGGCAGTTGCTGGCTGATT 1452
 Db 402 ValAspLeuMetSerIleGlyThrLeuLeuAlaTyrSerLeuValAlaAlaCysValLeu 421
 QY 1453 TATCTTCGATACAAA 1467
 Db 422 ValLeuArgTyrGln 426

RESULT 13

US-09-134-001C-5370

; Sequence 5370, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 5370

; LENGTH: 499

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5370

Alignment Scores:

Pred. No.: 7,68e-15 Length: 499
 Score: 224.50 Matches: 122
 Percent Similarity: 41.90% Conservative: 90
 Best Local Similarity: 24.11% Mismatches: 193
 Query Match: 5.45% Indels: 101
 DB: 4 Gaps: 26

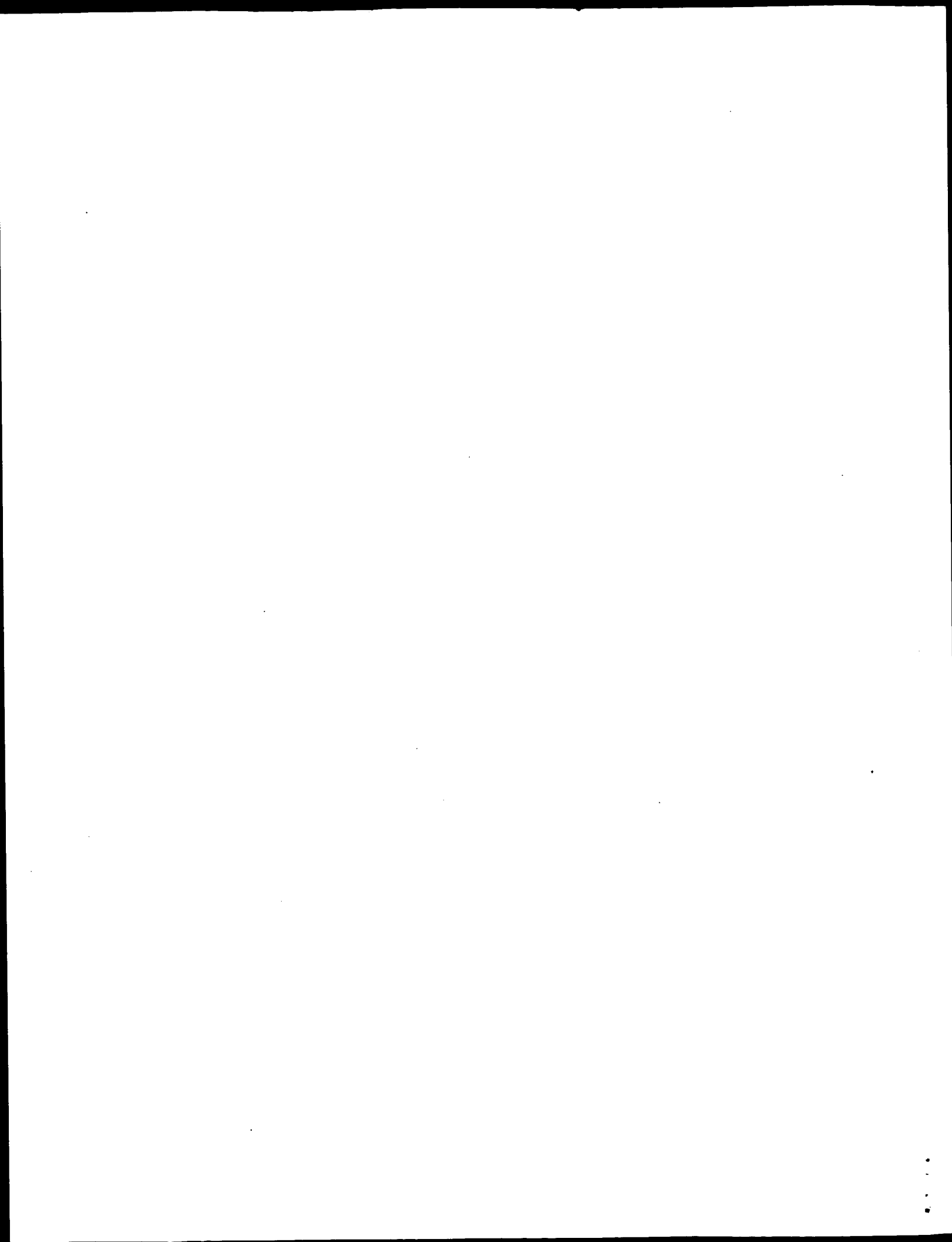
US-09-667-170A-440 (1-2239) x US-09-134-001C-5370 (1-499)

QY 379 TCATTATCATTTGGACCATCATTTGGACGAGGAATCTTCATCTCTCTAAGGGCGTCTC 438
 Db 27 AlaIleAlaTyrGlySerCysIleGlyTrpGlyAlaPheIleLeuProGlyAspTrpIle 46
 QY 439 CAGACACGGCGACGCGTGGCATGCTCTGACCATCTGACGGTCTGTTGGGTCCTGCTCA 498
 Db 47 LysGlnSerGlyProIleAlaSerIleGlyIle---ValIleGlyAlaLeuLeuMet 65
 QY 499 CTATTTGGAGCTTGTCTATCTCTGAATTTGGGAACAACATAAAGAACTTGGAGGTGCAT 558
 Db 66 IleLeuIleAlaValSerTyrGlyAlaLeuValGluLysPheProValSerGlyGlyAla 85
 QY 559 TACACATATATTGGAAGTCTCTTGGTCCATTACCATCTTTGTACGAGTCTGGGTGGAA 618
 Db 86 PheAlaPheSerPheLeuSerPheGlyArgTyrValSerPhePheSerSerTrpPhe--- 104
 QY 619 CTCTCATATACGCCCTGACGACTGCTGTGATATCCCTGGCATTTGGACGCTACATT 678
 Db 105 LeuThrPheGlyTyrValCysValValAlaLeuAsnAlaThrAlaPheSer----- 121
 QY 679 CTGGAACCATTTTATTCAATGTGAATCCCTGCACTTCCGATC-----AAGCTCATTT 732
 Db 122 -----LeuLeuIleLysPheLeuLeuProAsnValLeuAsnAsnGlyLysLeuTyr 138
 QY 733 ACAGCTGTGGGC-----ATAACT 750

Db 139 ThrValAlaGlyTrpAspValTyrIleThrGluIleValIleAlaThrValLeuLeuIle 158
 QY 751 GTAGTGTATGTCCTTAATAGCATGAGTGTGAGTGGAGCGCCGAGATTCAGATTTCCTTA 810
 Db 159 ValPheMetLeuIleThrIleArgGlyAlaSerValSerGlySerLeuGlnTyrTyr--- 177
 QY 811 ACCTTTTCCAAAGCTCACAGCAATTCATTAATATATAGTCCCTGGAGTTATGAGCTAATT 870
 Db 178 ---PheCys-----ValAlaMetValLeuValValAlaLeuMet----- 189
 QY 871 AAAGTCAAAAGCAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAGATTAATACGGGG 930
 Db 190 -----PheIleGlySerPhePheSerSerHisPheSerLeuSerHis 203
 QY 931 TTG---CCACTGGCT-----TTTTATTATGAATGTATGCA 963
 Db 204 LeuGluProLeuAlaSerValAspLysGlyTrpPheGlnSerIleIleMetIleValSer 223
 QY 964 TATGCTGGCTGGTTTACCTCAACTTT-----GTTACTCAAGAACTAGAA 1008
 Db 224 IleAlaProTrpAlaTyrValGlyPheAspAsnIleProGlnThrAlaGluGluPheAsn 243
 QY 1009 ---AACCTCAAAAACCATTCCTTGGCAATATATATATATATATATATATATATATAT 1065
 Db 244 PheSerProAsnLysThrPheLysLeu---IleValTyrSerLeuLeuAlaAlaSerLeu 262
 QY 1066 GGCTATGTGTGTGACAAATGTG-----GCCTACTTTACAGACCATTAAT 1107
 Db 263 ThrTyrValValMetLeuLeuTyrThrGlyTrpLeuSerThrGlnAlaThrSerLeuAsn 282
 QY 1108 GCTGAGGAGCTGCTCTTCAATGAGTGGCAGTGCAGTCTTCTGAGCGGCTACTGGGA 1167
 Db 283 Gly---AsnLeuTrpLeuThrGlyAlaVal-----ThrGlnAspAlaPheGly 297
 QY 1168 AATTTCTCATAGCAGTT---CCGATCTTTGTCCTCTCTCTCTCTCTCTCTCTCTCTCT 1224
 Db 298 PheIleGlyLeuAlaValAlaLeuAlaIleIleMetGlyIlePheThrGlyLeuAsn 317
 QY 1225 GGTGGTGTGTGTGCTCCAGGTTATTTCTATGTTGCGTCTCGAGAGGTCACCTTCCA 1284
 Db 318 GlyPheLeuMetSerSerArgLeuLeuPheSerMetGlyArgSerGlyIleMetPro 337
 QY 1285 GAAATCCTCTCCATGATTGTCGCGCAAGCACCTCTCTACAGCTGTATTTGTTTG 1344
 Db 338 ThrValPheSerLysLeuHisSerLysHisLysThrProTyrValAlaIleIlePheLeu 357
 QY 1345 CACCTTTGACAATGATAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
 Db 358 ValAlaValSerLeuIleAlaProTrpLeuGlyArgThrAlaLeuThrTrpIleValAsp 377
 QY 1381 CTCGACAGCTCTTTGAATTTCTCAGTTTGGCCAGTGGCTTTTATTGGGCTGGCAGTT 1440
 Db 378 MetSerSerThrGlyValSerIleAlaTyr-----PheIleThrCysLeuSer 393
 QY 1441 GCTGGGCTGATTTCTTCGATACAAATGCCAGATATGTCATCGTCTCT---TTCAGGTG 1497
 Db 394 AlaThrLysLeuPheSerPheAsnLysGlnSerAsnThrTyrAlaProValTyrLysIle 413
 QY 1498 CCAGTGTTCATCCACCTTTG-----TTTCTCTTCACATGCTCTCTCTCTCTCTCTCT 1545
 Db 414 PheGlyIleIleGlySerIleValSerPheValPheLeuCysLeuLeuIleProGly 433
 QY 1546 -----GCCCTTTCCCTCTATTTCGGACCATTTTACAGAGGATTTGGCTTCGTCATC 1596
 Db 434 SerProAlaAlaLeuSerIleProSer---TyrIleAlaLeuGlyIleTrpLeuValIle 452
 QY 1597 ACTCTACCTGGAGTCCCTCGGTATTTCTCTTTTATATATATATATATATATATATATAT 1656
 Db 453 GlyLeu-----IlePhePheIleIle-----ArgLeuProLysLeu 464
 QY 1657 TTAGAATAATCTCAGAGAAAATACACAGACATTA-----CAATAATA 1701
 Db 465 LysLysMetAsnAsnAspGluLeuSerArgLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 484

QY 1702 CTGGAAGTTGTACAGAA 1719
 |||||:|||||
 Db 485 LeuGluMetValHisGlu 490
 |||||:|||||
 RESULT 14
 US-09-134-001C-4475
 ; Sequence 4475, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 4475
 ; LENGTH: 494
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-4475
 Alignment Scores:
 Pred. No.: 2,56e-12 Length: 494
 Score: 200.50 Matches: 97
 Percent Similarity: 43.02% Conservative: 88
 Best Local Similarity: 22.56% Mismatches: 176
 Indels: 69
 Gaps: 20
 Query Match: 4
 DB:
 US-09-667-170A-440 (1-2239) x US-09-134-001C-4475 (1-494)
 QY 547 TCTGGAGTCAATACACATATATTTGGAGAGCTTTGGTCCATTACCAGCTTTTCTACGA 606
 |||||:|||||
 Db 8 SerGlyGlyMetSerAsnTyAlaGlnTyThrHisGlySerLeuLeuGlyPheIleAla 27
 |||||:|||||
 QY 607 GTC-----TGGTGGAACTCCTCATATAATAGCCCT-----GCAGCTACTGCTGTG 651
 |||||:|||||
 Db 28 AlaTrpAlaAsnTrpValSerLeuValThrIleIleProIleGluAlaValSerAlaVal 47
 |||||:|||||
 QY 652 ATATCCCTGGCATTGGAGCTACATT-----CTGGAACCATTT 690
 |||||:|||||
 Db 48 -----GlnTyMetSerSerTrpProTrpAspTrpAlaLysPro--- 60
 QY 691 TTTATTCAATGTGAATCCCTGAATCCGTAATCGCATCAAGCTCATTACAGCTGGGGCATAACT 750
 |||||:|||||
 Db 61 ---MetGlySerLeuMetLysAspGlySerIleSerThrTyGlyLeuIleAlaValTy 79
 |||||:|||||
 QY 751 GTAGTGATGCTCTAAATAGCANGAGTGTACGTGGAGCGCCGGATCCAGATTTCTTA 810
 |||||:|||||
 Db 80 IleIleIleAlaIlePheSerLeuLeuAsnTyTrpSer-----ValLysLeuLeuThr 97
 |||||:|||||
 QY 811 ACCTTTTGCAGCTCACACCAATTCGTATAATTATAGTCCCT-----GGAGTTATG 861
 |||||:|||||
 Db 98 SerPheThrSerLeuIleSerValPheLysIleGlyValProIleLeuThrIleIleMet 117
 |||||:|||||
 QY 862 CAGCTAATAAAGGT---CAAAAGCGACAACCTTAAGACGCCCTTTTCAGGAAGAGATTCA 918
 |||||:|||||
 Db 118 LeuLeuValSerGlyPheAspThrGlyAsnTyGlyHisSerIleGly----- 133
 |||||:|||||
 QY 919 AGTATTACGGGTGGCCACTGGCTTTTATATGA-----TyrGlySerAlaProIlePheAlaAlaThr 147
 |||||:|||||
 Db 134 -----ThrPheMetPro----- 147
 |||||:|||||
 QY 955 -----ATGTATGCATATGCTGGCTGCTTTTACTCAACTTGTCTACTGAA 999
 |||||:|||||
 Db 149 ThrThrSerGlyIleIlePheSerPheAsnAlaPheGlnThrIleIleAsnMetGlySer 167
 |||||:|||||

Search completed: April 16, 2003, 16:27:03
Job time : 35 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:10:40 ; Search time 65.5 Seconds

(without alignments)
5169.450 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 288829 seqs, 75613885 residues

Total number of hits satisfying chosen parameters: 577658

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09667170/runat_04042003_090919_20390/app_query.fasta_1.2375
-DB=Published Applications_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosom62
-TRANS=human40.cdi -LIST=45 -MODS=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09667170 -CGN1.1.15 -runat_04042003_090919_20390
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:
1: /cgn2.6/ptodata/1/pubpaa/US08_NEW_PUB_PEP.*
2: /cgn2.6/ptodata/1/pubpaa/PCT_NEW_PUB_PEP.*
3: /cgn2.6/ptodata/1/pubpaa/US06_NEW_PUB_PEP.*
4: /cgn2.6/ptodata/1/pubpaa/US06_PUBCOMB_PEP.*
5: /cgn2.6/ptodata/1/pubpaa/US07_NEW_PUB_PEP.*
6: /cgn2.6/ptodata/1/pubpaa/US07_PUBCOMB_PEP.*
7: /cgn2.6/ptodata/1/pubpaa/PCTUS_PUBCOMB_PEP.*
8: /cgn2.6/ptodata/1/pubpaa/US08_PUBCOMB_PEP.*
9: /cgn2.6/ptodata/1/pubpaa/US09_NEW_PUB_PEP.*
10: /cgn2.6/ptodata/1/pubpaa/US09_PUBCOMB_PEP.*
11: /cgn2.6/ptodata/1/pubpaa/US10_NEW_PUB_PEP.*
12: /cgn2.6/ptodata/1/pubpaa/US10_PUBCOMB_PEP.*
13: /cgn2.6/ptodata/1/pubpaa/US60_NEW_PUB_PEP.*
14: /cgn2.6/ptodata/1/pubpaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2519	61.1	501	9	US-10-163-866-48 Sequence 48, Appl
2	2519	61.1	501	9	US-10-163-866-49 Sequence 49, Appl
3	1210.5	29.4	507	9	US-10-163-866-54 Sequence 54, Appl
4	1206.5	29.3	524	9	US-10-163-866-37 Sequence 37, Appl

5	1198.5	29.1	507	9	US-10-163-866-38 Sequence 38, Appl
6	1198.5	29.1	507	9	US-10-163-866-39 Sequence 39, Appl
7	1156.5	28.1	515	9	US-10-163-866-40 Sequence 40, Appl
8	1131.3	27.5	511	9	US-10-163-866-41 Sequence 41, Appl
9	1131.3	27.5	511	9	US-10-163-866-42 Sequence 42, Appl
10	1101.5	26.7	535	9	US-10-163-866-43 Sequence 43, Appl
11	1101.5	26.7	535	9	US-10-163-866-44 Sequence 44, Appl
12	1075.5	26.1	487	9	US-10-163-866-46 Sequence 46, Appl
13	1067.5	25.9	487	9	US-10-163-866-45 Sequence 45, Appl
14	1048	25.4	517	9	US-09-815-923-16 Sequence 16, Appl
15	1039	25.2	523	9	US-10-163-866-47 Sequence 47, Appl
16	732	17.8	414	10	US-09-925-297-747 Sequence 747, Appl
17	513	12.5	97	9	US-09-854-133-586 Sequence 586, Appl
18	513	12.5	97	10	US-09-738-973-586 Sequence 586, Appl
19	412.5	10.0	180	9	US-10-163-866-50 Sequence 50, Appl
20	406.5	9.9	179	10	US-09-864-761-43216 Sequence 43216, A
21	322	7.8	456	10	US-09-815-242-4932 Sequence 4932, A
22	322	7.8	463	10	US-09-815-242-4932 Sequence 10662, A
23	295	7.2	619	10	US-09-741-153-4 Sequence 4, Appl
24	286.5	7.0	619	10	US-09-741-153-4 Sequence 2, Appl
25	262	6.4	470	10	US-09-815-242-12078 Sequence 12078, A
26	257.5	6.2	482	10	US-09-815-242-12941 Sequence 12941, A
27	233.5	5.7	463	10	US-09-815-242-13918 Sequence 13918, A
28	232	5.6	475	10	US-09-815-242-10052 Sequence 10052, A
29	230.5	5.6	453	9	US-09-738-626-4026 Sequence 4026, Ap
30	229	5.6	468	9	US-09-738-626-4777 Sequence 4777, Ap
31	225.5	5.5	149	9	US-10-043-487-376 Sequence 376, App
32	223.5	5.4	497	10	US-09-815-242-12626 Sequence 12626, A
33	219	5.3	469	10	US-09-815-242-5281 Sequence 5281, Ap
34	218.5	5.3	489	10	US-09-815-242-13932 Sequence 13932, A
35	216.5	5.3	489	10	US-09-815-242-10215 Sequence 10215, A
36	208.5	5.1	475	10	US-09-815-242-11503 Sequence 11503, A
37	207.5	5.0	519	9	US-09-895-913A-118 Sequence 118, App
38	207.5	5.0	519	9	US-09-815-242-11388 Sequence 11388, A
39	199.5	4.8	487	10	US-09-815-242-12036 Sequence 12036, A
40	197.5	4.8	459	10	US-09-815-242-12844 Sequence 12844, A
41	176	4.3	603	9	US-09-738-626-6946 Sequence 6946, Ap
42	174	4.2	603	10	US-09-801-368-144 Sequence 144, App
43	173	4.2	466	10	US-09-815-242-10265 Sequence 10265, A
44	173	4.2	475	10	US-09-815-242-5055 Sequence 5055, Ap
45	162	3.9	466	10	US-09-815-242-14048 Sequence 14048, A

ALIGNMENTS

RESULT 1
US-10-163-866-48
; Sequence 48, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-48

Alignment Scores:

Pred. No.: 1.47e-237 Length: 501
 Score: 2519.00 Matches: 498
 Percent Similarity: 96.51% Conservativity: 0
 Best Local Similarity: 96.51% Mismatches: 0
 Query Match: 61.14% Indels: 18
 DB: 9 Gaps: 1

US-09-667-170A-440 (1-2239) x US-10-163-866-48 (1-501)

```

QY 184 AAGCCTGTGTGTCACCATCTCCAAAGGAGGTACCTCCAGGAAATGTTAAACGGGAGG 243
Db 4 LysProValValSerThrIleSerLysGlyGlyTyrLeuGlnGlnValAsnGlyArg 23
QY 244 CTGCCTTCCTGGGCAACAAGGAGCCACCTGGGACAGGAGCCCTTTTCAGGAAGAGACGCC 303
Db 24 LeuProSerLeuGlyAsnLysGluProGlyGln----- 35
QY 304 TTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTGCAGCTCAGAGGAAAGTCACT 363
Db 36 -----GluLysValGlnLeuLysArgLysValThr 45
QY 364 TTACTGAGGGAGTCTCCATTATCATTTGGCACCATTATGGACGAGGAATCTTCATCTCT 423
Db 46 LeuLeuArgGlyValSerIleIleIleGlyThrIleIleGlyAlaGlyIlePheIleSer 65
QY 424 CCTAAGGCGTGTCTCCAGAACACGGGACGGCGTGCATGTCTCTGACCATCTGGACGGTG 483
Db 66 ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleTriThrVal 85
QY 484 TGTGGGTCCTGTCACTATTGTGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAAAG 543
Db 86 CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrIleLys 105
QY 544 AAATCTGGAGGTCAATACACATATATTTTGGAAGTCTTTTGGTCCATTACAGCTTTTGTA 603
Db 106 LysSerGlyGlyHisTyrThrTyrIleLeuGluValPheGlyProLeuProAlaPheVal 125
QY 604 CGAGTCTGGTGGAACTCCTCATATAGCCCTCCAGCTACTGCTGTGATATCCCTGGCA 663
Db 126 ArgValTrpValGluLeuLeuIleIleArgProAlaAlaThrAlaValIleSerLeuAla 145
QY 664 TTTGGACGCTACATCTCTGGAACCACTTTTATTCAATGTGAATCCCTGAACTTGCAGTC 723
Db 146 PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle 165
QY 724 AAGCTCATACAGCTGTGGGATAACTGTAGTGTGATGCTCTTAATAGCATGAGTCAGC 783
Db 166 LysLeuIleThrAlaValGlyIleThrValValMetValLeuAsnSerMetSerValSer 185
QY 784 TGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCAACAGCTCAACAGCGCTTT 903
Db 186 TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIle 205
QY 844 ATAGTCCCTGGAGTATGCACTAATTAAGGTCAACAGCAGAACTTTAAAGACGCGCTT 903
Db 206 IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPheLysAspAlaPhe 225
QY 904 TCAGGAAGAGATTCAGATTTACCGGTTGCCACTGCTTTTATATGGAATGATGCA 963
Db 226 SerGlyArgAspSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla 245
QY 964 TATCCTGGCTGGTTTACCTCAACTTTGTACTGAAGAAGTGAAGAAACCTGAAAAAAC 1023
Db 246 TyrAlaGlyTrpPheTyrLeuAsnPheValThrGluGluValGluAsnProGluLysThr 265
QY 1024 ATTCCTCTTCAATATGATATCATGTCATGTCATGTCACCATTTGGCTATGCTGACAAAT 1083
Db 266 IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn 285
QY 1084 GTGCGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTCTTCAATATGCTGAGTCAGTG 1143
Db 286 ValAlaTyrPheThrThrIleAsnAlaGluLeuLeuLeuSerAsnAlaValAlaVal 305
  
```

```

QY 1144 ACCTTTTCTGAGCGGTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTTGTCCCTC 1203
Db 306 ThrPheSerGluArgLeuLeuGlyAsnPheSerLeuAlaValProIlePheValAlaLeu 325
QY 1204 TCCTGCTTTGGCTCCATGACGGTGGTGTGTGCTGCTCCAGGTTATTTATGTTGCG 1263
Db 326 SerCysPheGlySerMetAsnGlyValPheAlaValSerArgLeuPheTyrValAla 345
QY 1264 TCTTCGAGAGGGTCACCTTCCAGAAATCTCTCATGATTTCATGTCGCGCAACACACCTCT 1323
Db 346 SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValArgLysHisThrPro 365
QY 1324 CTACACAGCTGTTATGTTTGTTCACCTTTTGACAATGATAATGCTCTTCTCTGGAGACTC 1383
Db 366 LeuProAlaValIleValLeuHisProLeuThrMetIleMetLeuPheSerGlyAspLeu 385
QY 1384 GACAGCTTTTGAATTTCTCAGTTTTCAGAGTGGCTTTTATTTGGCTGGCGAGTGGCT 1443
Db 386 AspSerLeuLeuAsnPheLeuSerPheAlaArgTrpLeuPheIleGlyLeuAlaValAla 405
QY 1444 GGGCTGATTATCTTCGATACAAATGCCAGATATGCATGCTCTTCAAGGTGCCACTG 1503
Db 406 GlyLeuIleTyrLeuArgTyrLysCysProAspMetHisArgProPheLysValProLeu 425
QY 1504 TTTATCCCGAGCTTTCTTTTCTTCATGCTCTTTCATGCTCTTCCCTCTATTTCG 1563
Db 426 PheIleProAlaLeuPheSerPheThrCysLeuPheMetValAlaLeuSerLeuTyrSer 445
QY 1564 GACCACTTATGATACAGGATTTGCTCTCATCACTCTGAGCTGAGTCCCTGGCTATTAT 1623
Db 446 AspProPheSerThrGlyIleGlyPheValIleThrLeuThrGlyValProAlaTyrTyr 465
QY 1624 CTCCTTTATATATGGGACAAAGAACCCAGGTGGTGTAGAATAATGTCAGAAAAATAACC 1683
Db 466 LeuPheIleIleTrpAspLysLysProArgTrpPheArgIleMetSerGluLysIleThr 485
QY 1684 AGAACATTACAAATAAATACTGGAAGTTGTACCAGAGAAAGATTA 1731
Db 486 ArgThrLeuGlnIleIleLeuGluValProGluGluAspLysLeu 501
  
```

RESULT 2

US-10-163-866-49
 ; Sequence 49, Application US/10163866
 ; Publication No. US20030027188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.
 ; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
 ; FILE REFERENCE: EX02-080C
 ; CURRENT APPLICATION NUMBER: US/10/163,866
 ; PRIOR FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: US 60/296,076
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/328,605
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/338,733
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: US 60/357,253
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR APPLICATION NUMBER: US 60/357,600
 ; PRIOR FILING DATE: 2002-02-15
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 49
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-163-866-49

Alignment Scores:
 Pred. No.: 1.47e-237 Length: 501
 Score: 2519.00 Matches: 498
 Percent Similarity: 96.51% Conservativity: 0

Best Local Similarity: 96.51% Mismatches: 0
Query Match: 61.14% Indels: 18
DB: 9 Gaps: 1

US-09-667-170A-440 (1-2239) x US-10-163-866-49 (1-501)

QY 184 AAGCCTCTGTGTCACCATCTCCAAAGGAGGTTACTCGAGGGAATCTTAACGGGAGG 243
DB 4 LysProValValSerThrIleSerLysGlyTyrLeuGlnGlyAsnValAsnGlyArg 23
QY 244 CTGCTTCCCTGGGACACAGGAGCCACCTGGGCGAGGCGCCTTTTCAGGAAGAGACGCC 303
DB 24 LeuProSerLeuGlyAsnLysGluProProGlyGln----- 35
QY 304 TTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTCAGCTGAAGAGAAAGTCACT 363
DB 36 -----GluLysValGlnLeuLysArgLysValThr 45
QY 364 TTACTGAGGGAGTCTCCATTAATATGATGACCATCATTTGGAGCAGGAATCTTCATCTCT 423
DB 46 LeuLeuArgGlyValSerIleIleIleGlyThrIleIleGlyAlaGlyIlePheIleSer 65
QY 424 CTTAAGGGGTGCTCCAGAACACGGGAGCGCTGGGCATGCTCTGACCATCTGGACGGTG 483
DB 66 ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleTrpThrVal 85
QY 484 TGTGGGGTCTGTCACTATTTGGAGCTTTGTCTTATGCTGAATGGGAACAACTATAAG 543
DB 86 CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrIleLys 105
QY 544 AAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTTGGTCCATTACCAGCTTTTGTGTA 603
DB 106 LysSerGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 603
QY 604 CGAGTCTGGGTGGAACTCCTCATATACGCCCTCGCAGTCTGCTGTGATATCCCTGGCA 663
DB 126 ArgValTrpValGluLeuLeuIleIleArgProAlaAlaThrAlaValIleSerLeuAla 145
QY 664 TTTGAGCGCTACATCTCGAACCATTCTTTTATCAATGTGAATCCCTGAATTCGGATC 723
DB 146 PheGlyArgTyrIleLeuGluProPheIleGlnCysGluIleProGluLeuAlaIle 165
QY 724 AAGCTATTACAGCTGTGGGATAACTGTAGTGTGCTTAATAGCTAGTGTGCTGACG 783
DB 166 LysLeuIleThrAlaValGlyIleThrValValMetValLeuAsnSerMetSerValSer 185
QY 784 TGGAGCGCCGATCCAGATTTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATAATT 843
DB 186 trpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIleIle 205
QY 844 ATAGTCCCTGGAGTTATGACAGCTAATTAAGGTCAACAGCGAGAACTTTAAAGACGCTTT 903
DB 206 IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPheLysAspAlaPhe 225
QY 904 TCAGAGAGAGATCAAGTATTACGCGGTGCCACTGGCTTTTATTATGGAATGTATGCA 963
DB 226 SerGlyArgAspSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla 245
QY 964 TATGCTGGTGGTTTACCTCACTTTGTTTGTGGAAGAGTAGAAACCCCTGAAAAACC 1023
DB 246 TyrAlaGlyTrpPheTyrLeuAsnPheValThrGluGluValGluAsnProGluLysThr 265
QY 1024 ATTCCCTTGCATATGATATCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1083
DB 266 IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn 285
QY 1084 GTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAATGCAAGTGGCAGTG 1143
DB 286 ValAlaTyrPheThrThrIleAsnAlaGluLeuLeuLeuSerAsnAlaValAlaVal 305
QY 1144 ACCTTTCTGAGCGGTACTGGGAATTTCTCATTAGCAGATTCGATCTTTGTCGCCCTC 1203
DB 306 ThrPheSerGluArgLeuLeuGlyAsnPheSerLeuAlaValProIlePheValAlaLeu 325

QY 1204 TCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTCTATGTTGG 1263
DB 326 SerCysPheGlySerMetAsnGlyGlyValPheAlaValSerArgLeuPheTyrValAla 345
QY 1264 TCTCGAGAGGCTCACCTTCCAGAAATCCTCTCCATCATTCATTCCTCGGCAAGCACACTCT 1323
DB 346 SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValArgLysHisThrPro 365
QY 1324 CTACCAAGCTGTTATGTTTGGACCCCTTTGACAAATGATAATGCTCTCTCTGGAGACCTC 1383
DB 366 LeuProAlaValIleValLeuHisProLeuThrMetIleMetLeuPheSerGlyAspLeu 385
QY 1384 GACAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCT 1443
DB 386 AspSerLeuLeuAsnPheLeuSerPheAlaArgTyrLeuPheIleGlyLeuAlaValAla 405
QY 1444 GGCTCATTTATCTTCGATACAAATGCCAGATATGCATGCTCTTCAAGTGGCAGCTG 1503
DB 406 GlyLeuIleTyrLeuArgTyrLysCysProAspMetHisArgProPheLysValProLeu 425
QY 1504 TTATCCCAAGCTTTGTTTTCCTTCACATGCCCTCTTCATGTTGGCTTCCCTCTATTTCG 1563
DB 426 PheIleProAlaLeuPheSerPheThrCysLeuPheMetValAlaLeuSerLeuTyrSer 445
QY 1564 GACCAATTTAGTACAGGATTTGGCTTCGTCTCATCTGCTGAGTCCCTGCGTATTAT 1623
DB 446 AspProPheSerThrGlyIleGlyPheValIleThrLeuThrGlyValProAlaTyrTyr 465
QY 1624 CTTCTTATATATGAGCAAGAACCCAGGTGGTGTAGATAATGCTCAGAGAAATAACC 1683
DB 466 LeuPheIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 1683
QY 1684 AGAATTTACAAATTAATCTGGAAGTTGTACCAAGAAAGATAAGTTA 1731
DB 486 ArgThrLeuGlnIleLeuLeuGluValProGluGluAspLysLeu 501

RESULT 3
US-10-163-866-54
; Sequence 54, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SIC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-54

Alignment Scores:
Pred. No.: 1.55e-109 Length: 507
Score: 1210.50 Matches: 232
Percent Similarity: 66.19% Conservative: 95
Best Local Similarity: 46.96% Mismatches: 164
Query Match: 29.38% Indels: 3
DB: 9 Gaps: 2

Thu Apr 17 09:48:24 2003

```

US-09-667-170A-440 (1-2239) x US-10-163-866-54 (1-507)
QY 247 CCTTCCCTGGGCAACAGAGCCACCTGGGAGGAGCGCTTTTTCAGGAGAGAGCGCCTTT 306
Db 14 ProAlaAlaGluGluGluGluGluAaArgGluLysMetLeuAlaAlaLysSerAlaAsp 33
QY 307 TCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAGAAAGTGCAGCTGAAGAGAAAGTCACTTTA 366
Db 34 GlySerAlaProAlaGluGluGly--GluGlyValThrLeuGlnArgAsnIleThrLeu 52
QY 367 CTGAGGGAGCTCCATTATCATTCATGGCACCACCATTCATGGAGGAGGAATCTTCATCTCTCT 426
Db 53 LeuAsnGlyValAlaIleValGlyThrIleIleGlySerGlyIlePheValThrPro 72
QY 427 AAGGCGCTGCTCAGAACACGCGGCGGCGCATGCTCTCTGACCATTCGACGGGTGTGT 486
Db 73 ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValValThrAlaAlaCys 92
QY 487 GGGTCTCTGCTCATTATTCAGTCTTGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546
Db 93 GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleSerLys 112
QY 547 TCTGAGGTCTATACATATATTTGGAAGTCTTTGGTCCATTTACCACTTTTGTACGA 606
Db 113 SerGlyGlyAspThrAlaThrMetLeuGluValThrGlySerLeuProAlaPheLeuLys 132
QY 607 GTCTGGGTGGAACCTCCTCATATACGCCCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
Db 133 LeuThrIleGluLeuLeuIleAlaArgProSerSerGlnTyrIleValAlaLeuValPhe 152
QY 667 GGACGCTACATCTGGAACCATTTTATTCATGTAATGCTGGAATCCTGCAATTCGATCAAG 726
Db 153 AlaThrThrLeuLeuLysProLeuPheProThrCysProValProGluAlaAlaLys 172
QY 727 CTCATTACAGCTGGGCGATCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
Db 173 LeuValAlaCysLeuCysValLeuLeuLeuAlaValAsnCysTyrSerValLysAla 192
QY 787 ACGCCCGGATCCAGATTTTCTTAACCTTTGCAAGCTCAGCAATTCGTGATAATATA 846
Db 193 AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAlaLeuIle 212
QY 847 GTCCCTGGAGTATGACCTAATTAAGGTCAAACGCGAGAACTTTAAAGACGCTTTTCA 906
Db 213 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 232
QY 907 -----GGAGAGATTCAGTATTACGGGTTGCCACTGGCTTTTATTTATGGAATGAT 960
Db 233 PheGlyGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuLysSerGlyLeuPhe 252
QY 961 GCATATGCTGCTGCTGCTTTTACCTCACTTGTCTGAGAGAGTGAAGAACCCCTGAAAAA 1020
Db 253 AlaThrGlyGlyThrAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
QY 1021 ACCATTCCCTGCAATATGATATCCATGGCCATTTGCACCATTTGCTATGCTGCTGACA 1080
Db 273 AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValThrValLeuThr 292
QY 1081 AATGTCGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATGCAGTGCCA 1140
Db 293 AsnLeuAlaThrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 312
QY 1141 GTGACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTTAGCAGTTCCTGCTTTGTTGCC 1200
Db 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerThrIleIleProValPheValGly 332
QY 1201 CTCTCTGCTTTGGCTCCATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPhePheVal 352
QY 1261 GCGTCTCGAGAGGCTCACCTTCCAGAAATCTCTCCATGATTCATGCTGCGGAGCACACT 1320
Db 353 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372

```

RESULT 4

```

US-10-163-866-37
; Sequence 37, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-37

```

```

Alignment Scores:
Score: 3.88e-109 Length: 524
Matches: 231
Conservative: 96
Percent Similarity: 66.19%
Best Local Similarity: 46.76%
Query Match: 29.28%
DB: 9
Gaps: 2

```

US-09-667-170A-440 (1-2239) x US-10-163-866-37 (1-524)

```

QY 247 CCTTCCCTGGGCAACAGAGCCACCTGGGAGGAGCGCTTTTTCAGGAGAGAGCGCCTTT 306
Db 31 ProAlaAlaGluGluLysLysLysGluAlaArgGluLysMetLeuAlaAlaLysSerAlaAsp 50
QY 307 TCAGGAAGAGAGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTCACTTTA 366

```


Db 51 GlySerAlaProAlaGluGluGly---GluGlyValThrLeuGlnArgAsnIleThrLeu 69
 QY 367 CTGAGGGAGCTCCATATATATGCGACCATATTTGGAGCAGGAATCTTCATCTCCT 426
 Db 70 LeuAsnGlyAlaAlaIleValGlyThrIleIleGlySerGlyIlePheValThrPro 89
 QY 427 AAGGCGCTGCTCCAGAACACCGCGGCGGTGGCATGCTCTGACCATCTGGACGGTGT 486
 Db 90 ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValThrAlaAlaCys 109
 QY 487 GGGCTCTGTCATATTTGGAGCTTTCTCTATGCTGAATGGGAACAACATATAAGAAA 546
 Db 110 GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrIleSerLys 129
 QY 547 TCTGAGGTCAATACACATATATTTGGAAGCTTTGGTCCATACCAGCTTTTGACGA 606
 Db 130 SerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAlaPheLeuLys 149
 QY 607 GTCTGGGTGAACCTCAATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTT 666
 Db 150 LeuTrpIleGluLeuLeuIleAlaArgProSerSerGlnTyrIleValAlaLeuValPhe 169
 QY 667 GGAGCTACATCTCGGAACCATTTTATTTCAATGTGAATCCTGAACCTGGCATCAAG 726
 Db 170 AlaThrTyrLeuLeuLysProLeuPheProThrCysProValProGluAlaAlaLys 189
 QY 727 CTCATTACAGCTGTGGCATAACTAGTGTGCTCTAAATAGCATGAGTGTACAGCTGG 786
 Db 190 LeuValAlaCysLeuLysValLeuLeuThrAlaValAsnCysTyrSerValLysAla 209
 QY 787 AGCGCCGGTCAAGATTTTCTTAACCTTTTGAAGCTCACAGCAATCTGATATATA 846
 Db 210 AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuIleIle 229
 QY 847 TTCCTGAGTACTGAGCTAAATTAAGCTCAACCCAGCAACTTTAAAGACCCCTTTTCA 906
 Db 230 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 249
 QY 907 -----GGAAGAGATTCAAGTATTACCGGTTGCCACTGGCTTTTATTATGGAATGAT 960
 Db 250 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 269
 QY 961 GCATATGCTGGCTGTTTACCTCACTTTGTTACTGAAGAGTAGAAACCCCTGAAAAA 1020
 Db 270 AlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 289
 QY 1021 ACCATTCCCTTGCATATATATCATGCGCATTTGTCACCATTTGGCTATGCTGACA 1080
 Db 290 AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValThrValLeuThr 309
 QY 1081 AATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTCAATGCAGTGCA 1140
 Db 310 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 329
 QY 1141 GTGACCTTTTCTGAGCGGTACTGGAATTTCTCATAGCATTTCCGATCTTTGTCCTCC 1200
 Db 330 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 349
 QY 1201 CTCTCCTGCTTTGGCTCCATGACGGTGTGTGTTGCTGCTCCAGGTATTTCTATCTT 1260
 Db 350 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPheVal 369
 QY 1261 GGTCTCGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCGAAGCACACT 1320
 Db 370 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 389
 QY 1321 CCTCTACAGCTGTTATTGTTTGGACCCCTTTGACCAATGATTAATGCTCTCTCTGGAGAC 1380
 Db 390 ProValProSerLeuLeuIlePheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 409
 QY 1381 CTCGACAGCTTTTGAATTTTCTAGTTTTCAGGTGGCTTTTATTATGGCTGGCAGTT 1440
 Db 410 IlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIle 429

QY 1441 GTTGGCTGATTTATCTTCGATAAATGCCAGATATCATCGTCCTTTCAAGTGCCA 1500
 Db 430 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 449
 QY 1501 CTGTTTCATCCAGCTTTGTTTCTTCACATGCCTCTTCATGTTGGCTTTTCCCTCTAT 1560
 Db 450 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 469
 QY 1561 TCGGACCCCATTTAGTACAGGATTGGCTTCGTCATCACTCTGACTGGAGTCCCTGGGTAT 1620
 Db 470 LysThrProValGluCysGlyIleGlyPheThrIleLeuSerGlyLeuProValTyr 489
 QY 1621 TATCTCTTTATTATATGGACAAAGAACCCAGGTGGTGTAGATAATATGTCAGAGAAAATA 1680
 Db 490 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr 509
 QY 1681 ACCAGAACATTACAATAATATGTAAGTTGTACCAAGAA 1722
 Db 510 ThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 523

RESULT 5

US-10-163-866-38
 ; Sequence 38, Application US/10163866
 ; Publication No. US20030027188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.
 ; TITLE OF INVENTION: SLIC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
 ; FILE REFERENCE: EX02-080C
 ; CURRENT APPLICATION NUMBER: US/10/163,866
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: US 60/296,076
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/328,605
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/338,733
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: US 60/357,253
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR APPLICATION NUMBER: US 60/357,600
 ; PRIOR FILING DATE: 2002-02-15
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 38
 ; LENGTH: 507
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-163-866-38

Alignment Scores:

Pred. No.: 2,32e-108 Length: 507
 Score: 1198.50 Matches: 231
 Percent Similarity: 65.79% Conservative: 94
 Best Local Similarity: 46.76% Mismatches: 166
 Query Match: 29.09% Indels: 3
 DB: 9 Gaps: 2

US-09-667-170A-440 (1-2239) x US-10-163-866-38 (1-507)

QY 247 CCTTCCCTGGGCAACAAAGACCCACCTGGGCGAGGACGCCCTTTTCAGGAAGAGAGCGCCTTT 306
 Db 14 ProValAlaGluGluLysGluGluAlaArgLysMetLeuAlaSerLysArgAlaAsp 33
 QY 307 TCAGGAACAGACCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTCACTTTA 366
 Db 34 GlyAlaAlaProAlaGlyGluGly---GluGlyValThrLeuGlnArgAsnIleThrLeu 52
 QY 367 CTGAGGAGCTCTCATTTATCATTTGGCACCATCATTTGGACAGCAATCTTCATCTCTCT 426
 Db 53 LeuAsnGlyValAlaIleIleValGlyAlaIleIleGlySerGlyIlePheValThrPro 72
 QY 427 AAGGCGTGTCTCCAGAACACAGCGGCGGTGGCATGCTCTGACCATCTGGAGCGGTGTGT 486

Db 113 SerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAlaPheLeuLys 132
QY 607 GTCTGGGTGGAACTCCCTCAATAATACGCCCTGAGCTACTGCTGTGATATCCCTGGCATTT 666
Db 133 LeuTrpIleGluLeuLeuIleAlaArgProSerSerGlnTyrIleValAlaLeuValPhe 152
QY 667 GGACGCTACATCTTGGGAACCATTTTATTCAATGTGAATCCCTGAACTTCGGATCAAG 726
Db 153 AlaAlaTyrLeuLeuLysProLeuPheProThrCysProValProGluGluAlaLys 172
QY 727 CTCATTACAGCTGTGGCATACTGTAGTCAGTGTCTAAATAGCATGAGTGTCTGAGCTGG 786
Db 173 LeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTyrSerValLysAla 192
QY 787 AGCCCGCGGATCCAGATTTCTTAAACCTTTGCAAGCTCACAGCAATCTGATTAATATA 846
Db 193 AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAlaLeuIleIle 212
QY 847 GTCCCTGGAGTTATGAGCTAATTAAGGTCAACGCGCAACTTTTAAAGAGCGCTTTTCA 906
Db 213 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 232
QY 907 -----GGAAGAGATTCAAGTATTAGCGGTTGCCACTGGCTTTTATATGGAATGAT 960
Db 233 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 252
QY 961 GCATATGCTGGCTGTTTACCTCAACTTTGTACTGAAGAGTAGAAGAACCCCTCAAAA 1020
Db 253 AlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
QY 1021 ACCATTCCCTTGCATATGATATATCCATGGCCATTGTACCATTTGGCTGTGACA 1080
Db 273 AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValTyrValLeuThr 292
QY 1081 AATGTGGCTACTTTACGACCAATTAATGCTGAGGAGTGTCTTCAATGTCAGTGGCA 1140
Db 293 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 312
QY 1141 GTGACCTTTTCTGACGGGTACTGGGAAATTTCTCATTAGCAGTTCGGATCTTTGTGTC 1200
Db 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 332
QY 1201 CTCTCTGCTTGTGGCTCCATGAAGGGTGTGTGTTGCTGTCTCCAGGTTATCTAGTT 1260
Db 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPheVal 352
QY 1261 GCGTCTCGAGAGGTACCTTTCCAGAAATCTCTCCATGATTATGTCGCCCAAGCACACT 1320
Db 353 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372
QY 1321 CCTCTACAGCTGTATTGTTTTCACCCCTTTCACAAATGATAATGCTCTCTCGAGAC 1380
Db 373 ProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 392
QY 1381 CTCGACAGCTTTTGAATTTCTCAGTTTTCGCCAGTGGCTTTTATGGCTGGCAGTT 1440
Db 393 IlePheSerValIleAsnPhePheSerPhePheAsnTrpLeuCysValAlaLeuAlaIle 412
QY 1441 GCTGGCTGATTTATCTTCGATACAAATGCCAGATATCATCGTCTCTTCAAGGTGCCA 1500
Db 413 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 432
QY 1501 CTGTTTATCCAGCTTGTGTTTCTTCCATGCTCTTCCATGCTGTGCTTCCCTCTAT 1560
Db 433 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 452
QY 1561 TCGGACCCATTTAGTACAGGATGCTGTCATCATCTGCTACTGAGTCCCTCGGTAT 1620
Db 453 LysThrProValGluCysGlyIleGlyPheThrIleIleLeuSerGlyLeuProValTyr 472
QY 1621 TATCTCTTTATATATGGGACAGAACCCAGGTGTTTAGAATATATGTCAGAGAAAAATA 1680
Db 473 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr 492

QY 1681 ACCAGAACATTAACAATAATACGAGTTGTACCAAGAA 1722
Db 493 ThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 506

RESULT 7

US-10-163-866-40
; Sequence 40, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-40

Alignment Scores:

Pred. No.:	3e-104	Length:	515
Score:	1156.50	Matches:	218
Percent Similarity:	69.16%	Conservative:	105
Best Local Similarity:	46.68%	Mismatches:	141
Query Match:	28.07%	Indels:	3
Db:	9	Gaps:	2

US-09-667-170A-440 (1-2239) x US-10-163-866-40 (1-515)

QY 334 GAGAAAGTGCAGCTCAAGAGAAAGTACITTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Db 37 GluThrMetGlnLeuLysLysGluIleSerLeuLeuAsnGlyValSerLeuValValGly 56
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCAAGGCGTGTCTCCAGAACACGGCAGC 453
Db 57 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuValHisThrAlaSer 76
QY 454 GTGGCAGTCTCTGACCATCTGGACGGTGTGTGGGTCTGTCTACTATTGAGGAGCTTTG 513
Db 77 TyrGlyMetSerLeuIleValTrpAlaIleGlyLeuPheSerValValGlyAlaLeu 96
QY 514 TCTTATGCTGAATTCGGAACAACCTATATAAGAAATCTGAGGTCATTACACATATATTTG 573
Db 97 CysTyrAlaGluLeuGlyThrThrIleThrLysSerGlyAlaSerTyrAlaTyrIleLeu 116
QY 574 GAATCTTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCTCATATATACGC 633
Db 117 GluAlaPheGlyGlyPheIleAlaPheIleArgLeuTrpValSerLeuLeuValValGlu 136
QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGAGCTACATTTCTTGAACCATTTT 693
Db 137 ProThrGlyGlnAlaIleIleAlaIleThrPheAlaAsnTyrIleIleGlnProSerPhe 156
QY 694 ATTCAATGTGAATCCCTGAACCTTGGCATCAAGCTCATTTACAGCTGTGGGCATACCTGTA 753
Db 157 ProSerCysAspProTyrLeuAlaCysArgLeuLeuAlaAlaCysIleCysLeu 176
QY 754 GTGATGCTCTTAATAGCATGAGTGTACGCTGGAGCGCCCGGATCCAGATTTTCTTAACC 813
Db 754 GTGATGCTCTTAATAGCATGAGTGTACGCTGGAGCGCCCGGATCCAGATTTTCTTAACC 813

FILE REFERENCE: EX02-080C
 CURRENT APPLICATION NUMBER: US/10/163,866
 CURRENT FILING DATE: 2002-06-05
 PRIOR APPLICATION NUMBER: US 60/296,076
 PRIOR FILING DATE: 2001-06-05
 PRIOR APPLICATION NUMBER: US 60/328,605
 PRIOR FILING DATE: 2001-10-10
 PRIOR APPLICATION NUMBER: US 60/338,733
 PRIOR FILING DATE: 2001-10-22
 PRIOR APPLICATION NUMBER: US 60/357,253
 PRIOR FILING DATE: 2002-02-15
 PRIOR APPLICATION NUMBER: US 60/357,600
 PRIOR FILING DATE: 2002-02-15
 NUMBER OF SEQ ID NOS: 54
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 41
 LENGTH: 511
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-163-866-41

Alignment Scores:
 Pred. No.: 8,36e-102 Length: 511
 Score: 1131.50 Matches: 214
 Percent Similarity: 68.00% Conservative: 109
 Best Local Similarity: 45.05% Mismatches: 149
 Query Match: 27.46% Indels: 3
 Gaps: 2
 DB:

US-09-667-170A-440 (1-2239) x US-10-163-866-41 (1-511)

QY	310	GGAAGAGACGGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTG	369
DB	21	GlyAspGlyAlaSerProGlyProGluGlnValLysLeuLysGluLeuSerLeuLeu	40
QY	370	AGGGAGAGTCCATTTATTCATTTGGCACCACATTTGGAGCAGCAATCTTCATCTCTCCTAAG	429
DB	41	AsnGlyValCysLeuLeuValGlyAsnMetIleGlySerGlyIlePheValSerProLys	60
QY	430	GGCGTGCTCCAGAACACAGCGGCGCATGTCTCTGCACCATCTGGACGGTGTGTGGG	489
DB	61	GlyValLeuLeuLeuSerAlaSerPheGlyLeuSerLeuValIleTrpAlaValGlyGly	80
QY	490	GTCTGTCTACTATTTGGAGCTTTGCTTATCTGAATTTGGGAACACATATAAGAAATCT	549
DB	81	LeuPheSerValPheGlyAlaLeuLeuGlyCysTyrAlaGluLeuGlyThrThrIleLysLysSer	100
QY	550	GGAGGTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACACAGCTTTGTACGAGTC	609
DB	101	GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu	120
QY	610	TGGGTGGAACTCCTCATATACGCCCTGCAGCTACTGCTGTATATCCCTGGCATTTTGA	669
DB	121	TrpThrSerLeuLeuIleLeuGluProThrSerGlnAlaIleIleAlaIleThrPheAla	140
QY	670	CGCTACATCTGGAACCATTTTATTCAATGTGAAATCCCTGAACTTCGCGATCAAGCTC	729
DB	141	AsnTyrMetValGlnProLeuPheProSerCysPheAlaProTyrAlaAlaSerArgLeu	160
QY	730	ATTACAGCTGTGGGCATAACTGTAGTGTGGTCTTAAATAGCATGAGTGTACGTGGAGC	789
DB	161	LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyrValLysTrpGly	180
QY	790	GCCGGATCCAGATTTTCITAACTTTTCAACCTTTCGACGCTCACAGCAATCTCTGATAATATGTC	849
DB	181	ThrLeuValGlnAspIlePheThrTyrAlaLysValLeuAlaLeuIleAlaValIleVal	200
QY	850	CCTGGAGTTATGCAGCTAATTAAGGTCAACAGCAGCAACTTTAAAGACGCCCTTTTCAGGA	909
DB	201	AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGlyGly	220
QY	910	AGAGATTCAGATATTACCGCGTTCGCCACTGGCTTTTATATGGAATGTATGATATGCT	969

RESULT 8
 US-10-163-866-41
 Sequence 41, Application US/10163866
 Publication No. US20030027188A1
 GENERAL INFORMATION:
 APPLICANT: EXELIXIS, INC.
 ADDRESS: 3000 WEST 10TH AVENUE, SUITE 100, DENVER, CO 80202
 ATTORNEY: KENNEDY KATZ & ASSOCIATES, P.C., 1000 17TH AVENUE, SUITE 1000, DENVER, CO 80202
 FILING OFFICE: USPTO, DENVER OFFICE, 1600 RAYMOND AVENUE, SUITE 1000, DENVER, CO 80202
 FILING DATE: 2002-06-05
 PRIORITY DATE: 2001-06-05
 PRIORITY APPLICATION: US 60/296,076
 PRIORITY APPLICATION NUMBER: US 60/296,076
 PRIORITY FILING DATE: 2001-06-05
 PRIORITY APPLICATION NUMBER: US 60/328,605
 PRIORITY FILING DATE: 2001-10-10
 PRIORITY APPLICATION NUMBER: US 60/338,733
 PRIORITY FILING DATE: 2001-10-22
 PRIORITY APPLICATION NUMBER: US 60/357,253
 PRIORITY FILING DATE: 2002-02-15
 PRIORITY APPLICATION NUMBER: US 60/357,600
 PRIORITY FILING DATE: 2002-02-15
 NUMBER OF SEQ ID NOS: 54
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 41
 LENGTH: 511
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-163-866-41

```

Db 221 SerSerPheAlaValGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer 240
QY 970 GGCTGGTTTACCTCAACTTTGTTACTGAGAGAGTAGAAACCTGAAAAACCATTCCTCC 1029
Db 241 GlyTrpAspThrLeuAsnTyrValThrGluGluIleLysAsnProGluArgAsnLeuPro 260
QY 1030 CTTGCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1089
Db 261 LeuSerIleGlyIleSerMetProIleValThrIleIleIleIleIleIleIleIleIle 280
QY 1090 TACTTTAGGACCAATATATATATATATATATATATATATATATATATATATATATAT 1149
Db 281 TyrTyrThrValLeuAspMetArgAspIleLeuAlaSerAlaValAlaValThrPhe 300
QY 1150 TCTGAGCGGCTACTGGGAAATTTCTCATATAGCAGTTCCAGTCTTTGTCCTCTCTGCG 1209
Db 301 AlaAspGlnIlePheGlyIlePheAsnTrpIleIleProLeuSerValAlaLeuSerCys 320
QY 1210 TTTGGCTCCATGAAAGCGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1269
Db 321 PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuPhePheValGlySerArg 340
QY 1270 GAGGCTCACCTCCAGAAATCTCTCCATGATTCATGTCGCGAAGCACACTCTCTACCA 1329
Db 341 GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro 360
QY 1330 GCTGTTATTTGTTGACCCCTTTGACAATGATAATCTCTCTCTGAGACCTCCACAGT 1389
Db 361 SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln 380
QY 1390 CTTTGGATTTCTCAGTTTGGCCAGTGGCTTTTATTTGGCTGGCAGTTGCTGGGCTG 1449
Db 381 LeuIleAsnTyrTyrSerPheSerTyrTrpPheValGlyLeuSerIleValGlyGln 400
QY 1450 ATTATCTTCGATACAAATGCCAGATATGATCGTCTCTTCAAGTGGCCACTGTTCTATC 1509
Db 401 LeuTyrLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerValPhePhe 420
QY 1510 CCAGCTTTGTTTCTTCACATGCCCTTCATGGTGGCTTTCCTCTCTATTCGACCCCA 1569
Db 421 ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr 440
QY 1570 TTTAGTACAGGATTTGGCTTCGTCATCACTCTGCTGAGTGGCCCTCGGTATATCTCT 1629
Db 441 IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIle 460
QY 1630 ATT-----ATATGGACAAAGAACCCAGGTGGTTTAGAATAATGTCAGAGAAATAACC 1683
Db 461 IleArgValProGluHisLysArgProLeuTyrLeuArgArgIleValGlySerAlaThr 480
QY 1684 AGAACATTACAATAATA---CTGGAAGTTGTACCAGAGAAGAT 1725
Db 481 ArgTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetAsp 495

```

RESULT 9

```

US-10-163-866-42
; Sequence 42, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600

```

```

; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-42

```

Alignment Scores:

```

Pred. No.: 8,36e-102 Length: 511
Score: 1131.50 Matches: 214
Percent Similarity: 68.00% Conservatives: 109
Best Local Similarity: 45.05% Mismatches: 149
Query Match: 27.46% Indels: 3
DB: 9 Gaps: 2

```

US-09-667-170A-440 (1-2239) x US-10-163-866-42 (1-511)

```

QY 310 GGAAGAGACGCTTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369
Db 21 GlyAspGlyAlaSerProGlyProGluGlnValLysLeuLysLysLysLysLysLysLys 40
QY 370 AGGGAGCTCTCCATTATTCATTCGACCCATTCATTCGACGAGGAGGAGGAGGAGGAGG 429
Db 41 AsnGlyValCysLeuIleValGlyAsnMetIleGlySerGlyIlePheValSerProLys 60
QY 430 GGGCTGTCTCCAGAACACGGGCGAGCTGGGCTGTCTCTGACCATCTGGACGCTGGTGGG 489
Db 61 GlyValLeuIleTyrSerAlaSerPheGlyLeuSerLeuValIleThrPalaValGlyGly 80
QY 490 GTCTGTCACTATTTGGAGCTTTGCTTATGCTGAATGGGAACAACACTATAAAGAAATCT 549
Db 81 LeuPheSerValPheGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleLysLysSer 100
QY 550 GGAGGTCATTAACATATATTTTGGAGGCTTTTGGTGGTCCATTCACGAGCTTTGTACGAGTC 609
Db 101 GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu 120
QY 610 TGGGTGGAGACTCTCATATAGCCCTCGAGCTACTGCTGTGATATCCCTGGCATTTTGGG 669
Db 121 TrpThrSerLeuLeuIleGluProThrSerGlnAlaIleIleAlaIleThrPheAla 140
QY 670 CGCTACATCTCGAACCATTTTATTCAAATGTGAATCCCTGAACTTGCATCAAGTC 729
Db 141 AsnTyrMetValGlnProLeuPheProSerCysPheAlaProTyrAlaAlaSerArgLeu 160
QY 730 ATTACAGCTGTGGGCATTAACCTAGTGTAGTGTCTCTAAATAGCATGAGTGCAGTGGAGC 789
Db 161 LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyrValLysTrpGly 180
QY 790 GCCCGGATCCAGATTTTCTTAACCTTTTTCAGAGCTCAGAGCAATCTGTGATAATATAGTC 849
Db 181 ThrLeuValGlnAspIlePheThrTyrAlaLysValLeuAlaLeuIleAlaValIleVal 200
QY 850 CTTGGAGTTATGCACTTAATTAAGGTCAAACGACAGAACTTTAAAGAGCGCTTTTCAGGA 909
Db 201 AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGlyGly 220
QY 910 AGAGATTCAAGTATTACCGGCTGTCCACTGGCTTTTATTATGGAATGTATGATCATGCT 969
Db 221 SerSerPheAlaValGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer 240
QY 970 GGCTGGTTTACCTCAACTTTGTTACTCAAGAAGTAGAAACCTGAAAAACCATTCCTCC 1029
Db 241 GlyTrpAspThrLeuAsnTyrValThrGluGluIleLysAsnProGluArgAsnLeuPro 260
QY 1030 CTTGCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1089
Db 261 LeuSerIleGlyIleSerMetProIleValThrIleIleIleIleIleIleIleIleIle 280
QY 1090 TACTTTAGGACCAATATATGCTGAGGAGCTGCTGCTTTCAAATGAGTGGCAGTGCACCTTT 1149

```

Db 281 TyrTyrThrValLeuAspMetArgAspIleLeuAlaSerAspAlaValAlaValThrPhe 300
 QY 1150 TCTGAGCGGTACTGGGAATTTCTCATTAGCAGTTCCTGATCTTGTTCCTCCTCTGC 1209
 Db 301 AlaAspGlnIlePheGlyIlePheAsnTyrIleIleProLeuSerValAlaLeuSerCys 320
 QY 1210 TTGGGCTCCATGAACGGTGTGTGTCTGCTCTCCAGGTATTCTATCTGCTGCTCCA 1269
 Db 321 PheGlyLeuLeuAsnAlaSerIleValAlaAlaSerArgLeuPheValGlySerArg 340
 QY 1270 GAGGTCACCTCCAGAAATCTCTCATGATTCATGTCGCAAGCACACCTCTCTACCA 1329
 Db 341 GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro 360
 QY 1330 GCTGTTATTTGTTGCCACCTTTGACAAATGATAATGCTCTCTGAGGACCTCGACAGT 1389
 Db 361 SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln 380
 QY 1390 CTTTGTGAATTCCTCAGCTTTTGGCAGGTGCTTTTATTTGGGCTGGCAGTGTCTGGCTG 1449
 Db 381 LeuIleAsnTyrTyrSerPheSerTyrTyrPhePheValGlyLeuSerIleValGlyGln 400
 QY 1450 ATTTATCTTCGATCAATGCCATGATGATGATGATGATGATGATGATGATGATGATGATG 1509
 Db 401 LeuTyrLeuArgTyrLysGluProAspArgProArgProLeuLysLeuSerValPhePhe 420
 QY 1510 CCAGCTTTGTTTCTTCATGCTCTTCATGCTTCCCTTCCCTTCTCTCTCTCTCTCTCTCT 1569
 Db 421 ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr 440
 QY 1570 TTTAGTACAGGATGCTTCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1629
 Db 441 IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIle 460
 QY 1630 ATT-----ATATGGGCAAGAACCCAGGTGGTTAGAAATATGTCAGAGAAATACCC 1683
 Db 461 IleArgValProGluHisLysArgProLeuTyrLeuArgIleValGlySerAlaThr 480
 QY 1684 AGACATTACAAATATA---CTGAAGTTCTACCAAGAGAT 1725
 Db 481 ArgTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetAsp 495

RESULT 10
 US-10-163-866-43
 ; Sequence 43, Application US/10163866
 ; Publication No. US20030027188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.
 ; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
 ; FILE REFERENCE: EX02-080C
 ; CURRENT APPLICATION NUMBER: US/10/163,866
 ; PRIOR FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: US 60/296,076
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/328,605
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/338,733
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: US 60/357,253
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR APPLICATION NUMBER: US 60/357,600
 ; PRIOR FILING DATE: 2002-02-15
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 43
 ; LENGTH: 535
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-163-866-43

Alignment Scores: 7,35e-99 Length: 535
 Pred. No.: 1101,50 Matches: 220
 Score:

Percent Similarity: 65.52% Conservative: 103
 Best Local Similarity: 44.62% Mismatches: 161
 Query Match: 26.74% Indels: 9
 DB: 4
 Gaps: 4
 US-09-667-170A-440 (1-2239) x US-10-163-866-43 (1-535)
 QY 262 AAGGAGCCACCTGGCAGGACGCCCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGACGCC 321
 Db 13 LysIysHisProGly-----GlyGlyGluSerAspAlaSerProGluAla 27
 QY 322 TTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTACCTTTTACTGAGGGAGTCTCC 381
 Db 28 GlySerGlyGlyGlyValAlaLeuLysGlyIleGlyLeuValSerAlaCysGly 47
 QY 382 ATTATCATTTGGCACCACATCTTGGGACGAGAAATCTTCATCTCTCTAAGGGCGTCTCCAG 441
 Db 48 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 67
 QY 442 AACACGGCGAGCGTGGCATGTCTCTGACCATCTGACGGTGTGTGGGGTCTCTGTCACTA 501
 Db 68 AsnAlaGlySerValGlyLeuAlaLeuIleValTyrIleValThrGlyPheIleThrVal 87
 QY 502 TTTGAGCTTTCTCTTATCTGTAATTTGGAAACAACTATAAAGAAATCTGGAGGTCAATAC 561
 Db 88 ValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr 107
 QY 562 ACATATATTTGGAAGTCTTCTGGTCCATTACCGAGTCTTGTACGAGTCTGGGTGGAACCT 621
 Db 108 SerTyrValLysAspIlePheGlyLeuAlaGlyPheLeuArgLeuTyrIleAlaVal 127
 QY 622 CTCATAATACGCCCTCCAGCTACTCTGTGTATATCCCTGGCATTTGGAGCTGACATCTCTG 681
 Db 128 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 147
 QY 682 GAACCATTTTTCATCAATCTGAAATCCCTGAACTTGGCATCAAGCTCAATCACTGCTGTG 741
 Db 148 GlnProLeuPheProThrCysPheProGluSerGlyLeuArgLeuLeuAlaIle 167
 QY 742 GGCATAACTGTAGTGTCTCTAAATAGCATGAGTGTGAGTGGAGCGCCGGATCCAG 801
 Db 168 CysLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTyrAlaThrArgValGln 187
 QY 802 ATTTCTTAACTTTTGGCAGCTCAGCAATCTCATATAATATAGTCCCTGAGTTAG 861
 Db 188 AspIlePheThrAlaGlyLysLeuLeuAlaLeuLeuIleIleIleMetGlyIleVal 207
 QY 862 CAGCTAATTAAGGTCAA-----ACGCAGAACTTTAAAGAGCCCTTTTTCAGGA---AGA 912
 Db 208 GlnIleCysLysGlyGlyGlyPheTyrPheTyrLeuGluProLysAsnAlaPheGluAsnPheGln 227
 QY 913 GATTCAAGTATTAACGGGTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972
 Db 228 GluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAlaTyrGlyGly 247
 QY 973 TGGTTTTCCTCACTTGTGTACTGAAGAGTAGAAGAACCTGAAACACCTTCCCTTCCCTT 1032
 Db 248 TyrAsnProLeuAsnTyrValThrGluGluLeuValAspProTyrLysAsnLeuProArg 267
 QY 1033 GCAATATGATATCCATGCGCATTTGTGTCACCATTTGGTGTATGCTGACAAATGGGCTTAC 1092
 Db 268 AlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsnValAlaTyr 287
 QY 1093 TTTAGCACCATTATGCTGAGGAGTCTGCTCTTCAATTCAGTTCAGTTCAGTTCAGTTCAGT 1152
 Db 288 ValThrAlaMetSerProGlnGluLeuLeuAlaSerAsnAlaValAlaValThrPheGly 307
 QY 1153 GAGGCGCTACTGGGAAATTTCTCATAGCAGTTCCCGATCTTTTGTGCCCTCTCTCTGCTT 1212
 Db 308 GluLysLeuLeuGlyValMetAlaTyrIleMetProIleSerValAlaLeuSerThrPhe 327
 QY 1213 GGTCCCATGAACGGTGTGTGTTGCTGCTCCAGGTATTCTATGTTGCTGCTCCAGAG 1272

Db 328 GlyGlyValAsnGlySerLeuPheThrSerSerArgLeuPheAlaGlyAlaArgGlu 347
 QY 1273 GGTCACTTCAGAAATCCTCTCCATCATTCATGTCGCAAGCACACTCTCTACAGGT 1332
 Db 348 GlyHisLeuProSerValLeuAlaMetIleHisValIysArgCysThrProIleProAla 367
 QY 1333 GTTATTGTTTGCACCCCTTTGACAAATGATATGCTCTCTCTGAGACCTCGACAGCTTT 1392
 Db 368 LeuLeuPheThrCysIleSerThrLeuLeuMetLeuValThrSerAspMetThrLeu 387
 QY 1393 TTGAATTCCTCAGCTTTTGCAGAGTGCTTTTATTTGGGTGCGAGTCTGCTGGCTGATT 1452
 Db 398 IleAsnThrValGlyPheIleAsnThrLeuPheThrGlyValThrValAlaGlyGlnIle 407
 QY 1453 TATCTTCGATACAAATGCCAGATATGATCATGCTCTTCAAGGTGCCACTGTTCATCCA 1512
 Db 408 ValLeuArgThrLysProAspIleProArgProIleLysIleAsnLeuLeuPhePro 427
 QY 1513 GCTTTGTTTCCCTTCACATGCTCTTCATGTTGCTTCCCTTCCTTCATTCGGACCCATT 1572
 Db 428 IleIleThrLeuLeuPheThrAlaPheLeuLeuValPheSerLeuThrProVal 447
 QY 1573 AGTACAGGATTGGCTTCGTCATCATCTCTGACGTGAGTCCCTCGGTATTCCTTATT 1632
 Db 448 ValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValThrPheLeuGlyVal 467
 QY 1633 ATATGGCAAGAAACCCAGGTGGTGTAGATAATGTCAGAGAAATAACCAAGACATTA 1692
 Db 468 TyrTrpGlnHisLysProLysCysPheSerAspPheIleGluLeuLeuThrLeuValSer 487
 QY 1693 CAATAATCTAGGAAGTTGTA---CCAGAAGAGATAAG 1728
 Db 488 GlnLysMetCysValValValTyrProGluValGluArg 500

RESULT 11

US-10-163-866-44
 ; Sequence 44, Application US/10163866
 ; Publication No. US20030027188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.
 ; TITLE OF INVENTION: SLG7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
 ; FILE REFERENCE: EX02-080C
 ; CURRENT APPLICATION NUMBER: US/10/163,866
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: US 60/296,076
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/328,605
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/338,733
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: US 60/357,253
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR APPLICATION NUMBER: US 60/357,600
 ; PRIOR FILING DATE: 2002-02-15
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 44
 ; LENGTH: 535
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-163-866-44

Alignment Scores:
 Pred. No.: 7,35e-99 Length: 535
 Score: 1101.50 Matches: 220
 Percent Similarity: 65.52% Conservative: 103
 Best Local Similarity: 44.62% Mismatches: 161
 Query Match: 26.74% Indels: 9
 DB: 9 Gaps: 4

US-09-667-170A-440 (1-2239) x US-10-163-866-44 (1-535)

QY 262 AAGGAGCCACCTGGGAGGAGCCCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGGCC 321

Db 13 LysLysHisProGly-----GlyGlyGluSerAspAlaGluAla 27
 QY 322 TTTTCAGGAAGAGAAAGTGCAGCTGAAGAGAGAAAGTCACTTTACTAGGGAGTCTCC 381
 Db 28 GlySerGlyGlyGlyValAlaLeuLysGlyGlyLeuValSerAlaCysGly 47
 QY 382 ATTATCATTTGCCACCATCATTTGGAGCAGGAATCTTCTATCTCTCTAAGCGGTGCTCCAG 441
 Db 48 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysSerGlyValLeuGlu 67
 QY 442 AACACGGCGCAGCTGGCGATCTCTGACCATCTGACGGGTGTGTGGGTGCTGTCATCTA 501
 Db 68 AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyPheIleThrVal 87
 QY 502 TTTGAGACCTTTGCTTATGCTGAATTTGGGAACAACATATAAGAAATCTCGAGGTCTATAC 561
 Db 88 ValGlyAlaLeuCysThrAlaGluLeuGlyValThrIleProLysSerGlyAspThr 107
 QY 562 ACATATATTTTGAAGTCTTTGTGCTCATTTACCAGCTTTTGTACGAGTCTGGGTGGAACCTC 621
 Db 108 SerThrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 127
 QY 622 CTCAATATACGCCCTGCAGCTACTGCTGTCATATCCCTGCCATTTGGAGCGCTACATCTG 681
 Db 128 LeuValIleThrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnThrValLeu 147
 QY 682 GAACCATTTTATCAATGTGAATCCCTGAATCTCGCATCAAGCTATTACAGCTGTG 741
 Db 148 GlnProLeuPheProThrCysPheProGluSerGlyLeuArgLeuAlaAlaIle 167
 QY 742 GGCATAACTGCTAGTGTGCTTAATACGATGTCAGTCAGTCAGTCGAGCGCCGAGTCCAG 801
 Db 168 CysLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 187
 QY 802 ATTTCTTAAACCTTTTGCAGCACTCACAGCAATCTGTGAATATATATATCCCTGGAGTATT 861
 Db 188 AspIlePheThrAlaGlyLysLeuAlaLeuAlaLeuIleIleMetGlyIleVal 207
 QY 862 CAGCTAATTAAGGTCAA-----ACGCAAGAACTTTAAAGACGCCCTTTTCAGGA---AGA 912
 Db 208 GlnIleCysLysGlyGlyThrPheTrpLeuGluProLysAsnAlaPheGluAsnPheGln 227
 QY 913 GATTCAAGTATTACGCGGTTGCCACTGGCTTTTATTATGGAATGATGATGCTGCTGGC 972
 Db 228 GluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAlaThrGlyGly 247
 QY 973 TGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAACCCCTGAAAAAACCATTTCCCTT 1032
 Db 248 TrpAsnPheLeuAsnThrValThrGluLeuValAspProThrLysAsnLeuProArg 267
 QY 1033 GCAATATGATATCATTCATGGCCATTTGTCACCATTTGGCTATGTGCTGACAAATGTGGCTAC 1092
 Db 268 AlaIlePheIleSerIleProLeuValThrPheValThrPheAlaAsnValAlaThr 287
 QY 1093 TTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAATGTCAGTGGCAGTACACCTTTCT 1152
 Db 288 ValThrAlaMetSerProGlnGluLeuLeuAlaSerAsnAlaValAlaValThrPheGly 307
 QY 1153 GAGCGGCTACTGGGAAATTTCTCATTTAGCTTCCGATCTTTGTCGCTCTCTCTCTCTCTCT 1212
 Db 308 GluLysLeuLeuGlyValMetAlaThrIleMetProIleSerValAlaLeuSerThrPhe 327
 QY 1213 GGCTCCATCAACGGTGTGTTGCTGCTCTCCAGGTTATTCTATGTTGGGTCTCGAGAG 1272
 Db 328 GlyGlyValAsnGlySerLeuPheThrSerArgLeuPheAlaGlyAlaArgGlu 347
 QY 1273 GGTCACTTCCAGAAATCTCTCCATGATTCTGTCGCAAGCACACTCTCTTACAGCT 1332
 Db 348 GlyHisLeuProSerValLeuAlaMetIleHisValIysArgCysThrProIleProAla 367
 QY 1333 GTTATTGTTTTCACCCCTTTGCATATGATATGCTCTCTCTGAGACCTTCGACAGCTCT 1392

Db 368 LeuLeuPheThrCysIleSerThrLeuLeuMetLeuValThrSerAspMetTyrThrLeu 387
Qy 1393 TTGAATTCCTCAGTTTTCAGGCTGGCTTTTATTGGCTGGCAGTTCCTGGCTGATT 1452
Db 388 IleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAlaGlyGlnIle 407
Qy 1453 TATCTCGATACAAATCCAGATATGATCGCTCTTTCAGGTGGCCACTGTTTCATCCCA 1512
Db 408 ValLeuArgTrpLysLysProAspIleProArgProIleLysIleAsnLeuLeuPhePro 427
Qy 1513 GCTTTGTTTCTTCACATGCCCTCTTCATGGTTGCCCTTTCCTCTATTCGGACCCATT 1572
Db 428 IleIleTyrLeuLeuPheThrAlaPheLeuLeuValPheSerLeuTrpSerGluProVal 447
Qy 1573 AGTACAGGATGGCTTCCTGCTATCCTGCTGAGTCCCTCGATATATCTCTTTATT 1632
Db 448 ValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPheLeuGlyVal 467
Qy 1633 ATATGGGCAAGAACCCAGGTGCTTTAGATATATGTCAGAGAAATAACACAGAACATTA 1692
Db 468 TyrTrpGlnHisLysProLysCysPheSerAspPheIleGluLeuLeuThrLeuValSer 487
Qy 1693 CAATAATACTAGGAAGTTGTA---CCAGAAGAGATAAG 1728
Db 488 GlnLysMetCysValValValTyrProGluValGluArg 500

RESULT 12
US-10-163-866-46
; Sequence 46, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLIC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-866-46

Alignment Scores:
Pred. No.: 2,47e-96 Length: 487
Score: 1075.50 Matches: 207
Percent Similarity: 66.16% Conservative: 98
Best Local Similarity: 44.90% Mismatches: 153
Query Match: 26.10% Indels: 3
DB: 9 Gaps: 2

US-09-667-170A-440 (1-2239) x US-10-163-866-46 (1-487)

Qy 346 CTGAAGAGAAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATTATGGA 405
Db 26 LeuGlnLysLeuGlyLeuIleSerGlyIleSerIleValGlyThrIleIleGly 45
Qy 406 GCAGGAATCTTCATCTCTCTAAGGCGTGCTCCAGAACACGGGCGGCGGATGCT 465
Db 46 SerGlyIlePheValSerProLysSerValLeuSerAsnThrGluAlaValGlyProCys 65
Qy 466 CTGACCATCTGGACGGGTGTGGGGTCTGCTACATATTTGGAGCTTTGCTTTATGCTGAA 525

Db 66 LeuIleIleTyrAlaAlaCysGlyValLeuAlaThrLeuGlyAlaLeuCysPheAlaGlu 85
Qy 526 TTGGGAACAACATATAAAGAAATCTGGAGCTCATTTACACATATATTTTGGGAAGTCTTTGGT 585
Db 86 LeuGlyThrMetIleThrLysSerGlyGlyGlyTyrProTyrLeuMetGluAlaTyrGly 105
Qy 586 CCATTACAGCTTTTGTACAGTCTGGGTGGAACCTCTCTCATATAATACGCCCTGACGCTACT 645
Db 106 ProIleProAlaTyrLeuPheSerTrpAlaSerLeuIleValIleLysProThrSerPhe 125
Qy 646 GCTGTATATCCCTGGCATTTGGAGCTGATCTGCGAACCACTTTTATTCATATGTA 705
Db 126 AlaIleIleCysLeuSerPheSerGluTyrValCysAlaProPheTyrValGlyCysLys 145
Qy 706 ATCCCTGAACCTGCGCATCAAGCTCATACAGCTGTGGCATAAATAGTGTAGTGTGCTCTA 765
Db 146 ProProGlnIleValValLysCysLeuAlaAlaAlaIleLeuPheIleSerThrVal 165
Qy 766 AATAGCATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGCAGCTC 825
Db 166 AsnSerLeuSerValArgLeuGlySerTyrValGlnAsnIlePheThrAlaAlaLysLeu 185
Qy 826 ACAGCAATCTGATATATATAGTCTCCCTGGAGTTATGCAGCTAATTAAGGTCACACGAG 885
Db 186 ValIleValAlaIleIleIleSerGlyLeuValLeuAlaGlnGlyAsnThrLys 205
Qy 886 AACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTACGCGGTTCGCACTGGCTTTT 945
Db 206 AsnPheAspAsnSerPheGluGlyAlaGlnLeuSerValGlyAlaIleSerLeuAlaPhe 225
Qy 946 TATATGGAATGATGATATGCTGGCTGGTGTGTTTACCTCACTTTGTTTACTGGAAGAATA 1005
Db 226 TyrAsnGlyLeuTrpAlaTyrAspGlyTyrPasnGlnLeuAsnTyrIleThrGluGluLeu 245
Qy 1006 GAAACCCCTGAAACACCATTCCTTGCATATGATATATCATCCATGCCATTTCTCATTAG 1065
Db 246 ArgAsnProTyrArgAsnLeuProLeuAlaIleIleIleGlyLeuProLeuValThrAla 265
Qy 1066 GGCTATGCTGTCACAAATGTGGCTACTTACGACCATTAATGCTGAGGAGCTGTGCTT 1125
Db 266 CystyrIleLeuMetAsnValSerTyrPheThrValMetThrAlaThrGluLeuGln 285
Qy 1126 TCAATGTCAGTGGCAGTGGCTTTCTGAGGCTACTCTGGAAATTTCTCTATTAGCAGTT 1185
Db 286 SerGlnAlaValAlaValThrPheGlyAspArgValLeuTyrProAlaSerIleVal 305
Qy 1186 CCGATCTTTGTCCTCTCTCTGCTTGGCTCCATGAACGGTGTGTGTTGCTGCTCTCC 1245
Db 306 ProLeuPheValAlaPheSerThrIleGlyAlaAlaAsnGlyThrCysPheThrAlaGly 325
Qy 1246 AGTTTATCTATGTTGGTCTCGAGGGTCACTTCCAGAAATCTCTCCATGATCAT 1305
Db 326 ArgLeuIleTyrValAlaGlyArgGlyHisMetLeuLysValLeuSerTyrIleSer 345
Qy 1306 GTCGCAAGCACACTCTCTACCACTGTTTATTGTTTGGCACCCTTTGCAACATGATAATG 1365
Db 346 ValArgLeuThrProAlaProAlaIleIlePheTyrGlyIleAlaThrIleTyr 365
Qy 1366 CTCTTCTCTGGAGACCTGACAGTCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTT 1425
Db 366 IleIleProGlyAspIleAsnSerLeuValAsnTyrPheSerPheAlaIleIlePhe 395
Qy 1426 ATTGGGCTGGCAGTGTGGGCTGATTATCTTTCATACAAATGCCAGATATGCAATCGT 1485
Db 386 TyrGlyLeuThrIleLeuGlyLeuIleValMetArgPheThrArgLysGluLeuArg 405
Qy 1486 CTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCATCATGCTCTCTCATGCTT 1545
Db 406 ProIleLysValProValValIleProValLeuMetThrLeuIleSerValPheLeuVal 425
Qy 1546 GCCCTTCCCTCTATTCGGACCCCATTTAGTACAGGGATTGCTTCGTCATC---ACTCTG 1602

Db 426 LeuAlaProIleIleSerLysProThrTrpGluTyrCysValLeuPheIleLeu 445
 QY 1603 ACTGGAGCCCTCGCTATTATCTCTTATATATGGACAGAAACCCAGGTGGTTAGA 1662
 Db 446 SerGlyLeuLeuPheTyrPheLeuPheValHisTyr-----LysPheGlyTrpAlaGln 463
 QY 1663 ATAATGTCAGAGAAATAACAGAACATTAACAATAATACTGGAAGTTCTGACCAAGAA 1722
 Db 464 LysIleSerLysProIleThrMetHisLeuGlnMetLeuMetGluValValProGlu 483
 QY 1723 GAT 1725
 Db 484 Glu 484

RESULT 13
 US-10-163-866-45
 ; Sequence 45, Application US/10163866
 ; Publication No. US20030027188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.
 ; TITLE OF INVENTION: SLIC7S AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
 ; FILE REFERENCE: EX02-080C
 ; CURRENT APPLICATION NUMBER: US/10/163,866
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: US 60/296,076
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/328,605
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/338,733
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: US 60/357,253
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR APPLICATION NUMBER: US 60/357,600
 ; PRIOR FILING DATE: 2002-02-15
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 45
 ; LENGTH: 487
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-163-866-45

Alignment Scores:
 Pred. No.: 1 5e-95 Length: 487
 Score: 1067.50 Matches: 206
 Percent Similarity: 65.94% Conservative: 98
 Best Local Similarity: 44.69% Mismatches: 154
 Query Match: 25.91% Indels: 3
 Db: 9 Gaps: 2

US-09-667-170A-440 (1-2239) x US-10-163-866-45 (1-487)

QY 346 CTGAAGAGNAAGTCACTTTACTAGGGAGTCTCCATTATCATTTGGCAGCATCATGGA 405
 Db 26 LeuGlnLysGluLeuGlyLeuIleSerGlyIleSerIleValGlyThrIleGly 45
 QY 406 GCAGGAATCTCATCTCTCAAGGCGCTCCAGAACACGGCAGCGGCGCATGTCT 465
 Db 46 SerGlyIlePheValSerSerLysSerValLeuSerAsnThrGluAlaValGlyProCys 65
 QY 466 CTCACCATCTGCACGGTGTGGGTCTGTCACTATTTGGAGCTTTGCTTATGCTGAA 525
 Db 66 LeuIleIleTrpAlaAlaCysGlyValLeuAlaThrLeuGlyAlaLeuCysPheAlaGlu 85
 QY 526 TTGGGAACAACATATAAGAAATCTGGAGGTCATTCACATATATTTTGAAGCTTTGGT 585
 Db 86 LeuGlyThrMetIleThrLysSerGlyGlyGlyTrpProThrLeuMetGluAlaThrGly 105
 QY 586 CCATTACAGCTTTTGTACGAGTCTGGGTGGAACTCTCATATACGCCCTCCACTACT 645
 Db 106 ProIleProAlaThrLeuPheSerTrpAlaSerLeuIleValIleLysProThrPhe 125
 QY 646 GCTGTGATATCCCTGGCATTTGGACCGCTACATCTCGGAACCATTTTATTCAATGTGAA 705

Db 126 AlaIleIleCysLeuSerPheSerGluTyrValCysAlaProPheTyrValGlyCysLys 145
 QY 706 ATCCCTGAACCTTCCGATCAAGCTCATTACAGCTGTGGGCATCACTGTAGTGATGGCCCTA 765
 Db 146 ProProGlnIleValValLysCysLeuAlaAlaIleLeuPheIleSerThrVal 165
 QY 766 AATACATGAGTGTGCTGAGCGCCCGGATCCAGATTTTCTTAACCTTTTCCAGCTC 825
 Db 166 AsnSerLeuSerValArgLeuGlySerTyrValGlnAsnIlePheThrAlaAlaLysLeu 185
 QY 826 ACAGCAATCTCGAATAATTATAGTCCCTGAGTTTACAGCTAATAAAGTCAACAGCGAG 885
 Db 186 ValIleValAlaIleIleIleSerGlyLeuValLeuAlaGlnGlyAsnThrLys 205
 QY 886 AACTTTAAGAGCGCTTTTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTT 945
 Db 206 AsnPheAspAsnSerPheGluGlyAlaGlnLeuSerValGlyAlaIleSerLeuAlaPhe 225
 QY 946 TATTATGGAATGTATGTCATATGCTGGTGGTGTTCCTCACTCACTTGTCTTACTGAAGA 1005
 Db 226 TyrAsnGlyLeuTrpAlaLysArgGlyTrpAsnGlnLeuAsnTyrIleThrGluGluLeu 245
 QY 1006 GAAACCCCTGAAACCAATTCCTCCCTGCAATATGATATCATCCATGCCATTGTCCACAT 1065
 Db 246 ArgAsnProTyrArgAsnLeuProLeuAlaIleIleGlyIleProLeuValThrAla 265
 QY 1066 GGCTATGTCTGACAAATGTGGCTACTTACGACCATTAATGCTGAGGAGCTGCTGCTT 1125
 Db 266 CysTyrIleLeuMetAsnValSerTyrPheThrValMetThrAlaThrGluLeuGln 285
 QY 1126 TCAATGTCAGTGGCAGTACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTT 1185
 Db 286 SerGlnAlaValAlaValThrPheGlyAspArgValLeuTyrProAlaSerTrpIleVal 305
 QY 1186 CCGATCTTTGTTGGCTCTCTCTGCTTGGCTCCATGAACGGTGGTGTGTGTCTCTCC 1245
 Db 306 ProLeuPheValAlaPheSerThrIleGlyAlaAlaAsnGlyThrCysPheThrAlaGly 325
 QY 1246 AGTTATCTATGTTGGCTCTCGAGAGGTCACCTTCCAGAAATCTCTCCATGATTCAT 1305
 Db 326 ArgLeuIleTyrValAlaGlyArgGluGlyHisMetLeuLysValLeuSerTyrIleSer 345
 QY 1306 GTCCGCAAGACACCTCTCTACAGCTGTATTGTTTTCACCTTTGACAAATGATAAG 1365
 Db 346 ValArgArgLeuThrProAlaProAlaIleIlePheTyrGlyIleAlaThrIleTyr 365
 QY 1366 CTCCTCTCTGAGACCTCGACAGTCTTTTCAATTTCTCAGTTTCCAGGTGGCTTTT 1425
 Db 366 IleIleProGlyAspIleAsnSerLeuValAsnTyrPheSerPheAlaAlaTrpLeuPhe 385
 QY 1426 ATTTGGCTGGCAGTGTGGCTGATTATCTTCGATACAAATGCCAGATATGATCGT 1485
 Db 386 TyrGlyLeuThrIleLeuGlyLeuIleValMetArgPheThrArgLysGluLeuGluArg 405
 QY 1486 CTTTCAAGTGGCAGCTGTTCATCCAGCTGTGTTTCTTCTTCCATGCTCTCATGGT 1545
 Db 406 ProIleLysValProValIleProValLeuMetThrLeuIleSerValPheLeuVal 425
 QY 1546 GCCTTTCTCTCTATTCGACCCCATTTTGTACAGGATTTGGCTTCCTCATC---ACCTG 1602
 Db 426 LeuAlaProIleIleSerLysProThrTrpGluTyrLeuTyrCysValLeuPheIleLeu 445
 QY 1603 ACTGGAGCCCTCGGTATTATCTCTTTTATATATGGGACAGAAACCCAGGTGGTTAGA 1662
 Db 446 SerGlyLeuLeuPheTyrPheLeuPheValHisTyr-----LysPheGlyTrpAlaGln 463
 QY 1663 ATAATCTCAGAGAAATACAGAACATTAACAATAATACTGGAAGTTGTACCAAGAA 1722
 Db 464 LysIleSerLysProIleThrMetHisLeuGlnMetLeuMetGluValValProGlu 483
 QY 1723 GAT 1725
 ; ; ;

484 Glu 484
 RESULT 14
 US-09-815-923-16
 ; Sequence 16, Application US/09815923
 ; Publication No. US20020197644A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilli, Sarjeet S.
 ; APPLICANT: Ross, Linda S.
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Use of Insect Cell Membrane Transporters as No. US20020197644A1
 ; TITLE OF INVENTION: Target Sites for Insecticides
 ; FILE REFERENCE: 023070-093800US
 ; CURRENT APPLICATION NUMBER: US/09/815,923
 ; CURRENT FILING DATE: 2001-03-23
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 16
 ; LENGTH: 517
 ; TYPE: PRT
 ; ORGANISM: Aedes aegypti
 ; FEATURE:
 ; OTHER INFORMATION: LAT transporter
 US-09-815-923-16
 Alignment Scores: Length: 517
 Pred. No.: 1,24e-93 Matches: 199
 Score: 1048.00
 Percent Similarity: 64.36% Conservative: 117
 Best Local Similarity: 40.53% Mismatches: 169
 Indels: 6
 Gaps: 3
 Query Match: 9
 DB:
 US-09-667-170A-440 (1-2239) x US-09-815-923-16 (1-517)
 QY 271 CTGGGAGGAGGCGCTTT-----TCAGGAAGAGAGCGCTTTTCAGGAAGA--- 315
 DB 23 ProGlyThrGluSerLeuProProIleGluAsnGlyAlaIleAlaGlySerGlyThrIle 42
 QY 316 GACGCGCTTTTCAGGAAGAGAGAAAGTCGACCTCGAAGAGGAAAGTCACCTTACTGAGGGA 375
 DB 43 AspGlyGlyGlyGlyAsnGluArgValIysMetIysGluLeuGlyLeuGluGly 62
 QY 376 GTCCTCATATCATTCATGGACCATCATTCGAGCAGGAATCTCATCTCTCTAAGGCGGTG 435
 DB 63 ValAlaIleIleGluGlyIleIlePheGlySerGlyIlePheIleSerProIysGlyVal 82
 QY 436 CTCCGAACACGCGGAGCGGTGGCGATGCTCTGACCATCTGGACGGGTGCTGGGGCTCG 495
 DB 83 LeuGlnGluValGlyAlaValGlyThrSerLeuValIleIleIleIleIleIleIle 102
 QY 496 TCACATTTGGAGCTTCTCTATCTCAATTTGGGAACAACATATAAAGAAATCTGGAGGT 555
 DB 103 SerMetIleGlyAlaLeuGlyCysTyrAlaGluLeuGlyThrAlaIleProLysSerGlyGly 122
 QY 556 CATTACACATATATTTTGGAGCTTTTGGTCTCATTCACCATTCAGGCTTTTGGACGAGTCTGGGTG 615
 DB 123 AspTyrAlaTyrIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 142
 QY 616 GAATCTCTATATACGCGCTCGACGTACTGCTGTGTATATCCCTGGCATTTTGGACGCTAC 675
 DB 143 AlaThrValIleIlePheValProSerThrAsnAlaIleMetGlyLeuThrPheAlaSerTyr 162
 QY 676 ATTCGTGAACCATTTTATTCATAA---TGTGAATCCCTGAACTTCGATCGACCTCAT 732
 DB 163 ValPheGlnProLeuPheAlaAlaGlyCysSerValProThrIleGlyLeuGlnLeuPhe 182
 QY 733 ACACTGTGGGCATACCTAGTGTAGTGGTCTCAATATGATGAGTGTCTAGCTGGAGCGCC 792
 DB 183 AlaAlaValThrIleCysLeuLeuThrTyrIleAsnAlaTyrAspValargValThrThr 202
 QY 703 CGGATCCAGATTTTCTTAACCTTTTGAAGCTTTGCAAGCTCACAGCAATCTCGATATATATAGTCCCT 852

; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/328,605
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/338,733
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: US 60/357,253
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR APPLICATION NUMBER: US 60/357,600
 ; PRIOR FILING DATE: 2002-02-15
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 47
 ; LENGTH: 523
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-163-866-47

Alignment Scores:
 Pred. No.: 9,48e-93 Length: 523
 Score: 1039.00 Matches: 211
 Percent Similarity: 64.40% Conservative: 102
 Best Local Similarity: 43.42% Mismatches: 157
 Query Match: 25.22% Indels: 16
 DB: Gaps: 5

US-09-667-170A-440 (1-2239) x US-10-163-866-47 (1-523)

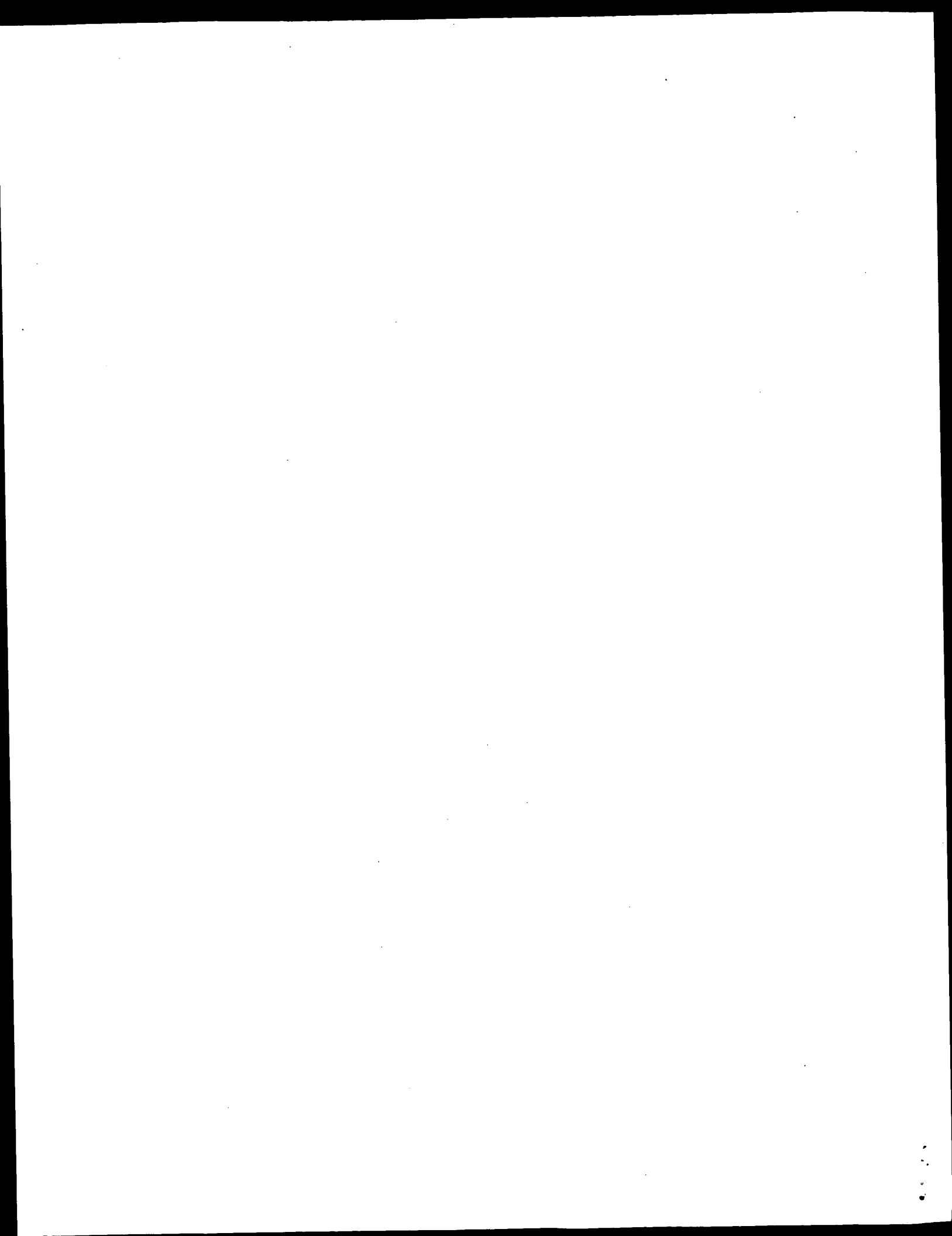
```

QY 334 GAGAAAGTGCAGTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTCATGGC 393
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 32  GluArgValAlaLeuLysGlyGluLeuLeuSerAlaCysThrIleIleIleGly 51
QY 394 ACCATCATTTGGACGAGATCTCATCTCTCTAAAGGGCGTCTCCAGAACACGGGACG 453
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 52  AsnIleIleGlySerGlyIlePheIleSerProLysGlyValLeuGluHisSerGlySer 71
QY 454 GTGGGATGCTCTGTACCATCTGGAGGCTGTGGGCTGCTGCTCACTATTGGAGCTTTG 513
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 72  ValGlyLeuAlaLeuPheValTrpValLeuGlyGlyValThrAlaLeuGlySerLeu 91
QY 514 TCTATGCTGAATTGGGAACAACATAAAGAAATCTGGAGGTCACTTACACATATATTTG 573
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 92  CysTyraAlaGluGlyValAlaIleProLysSerGlyGlyAspTyraAlaTyraValThr 111
QY 574 GAAGTCTTTGGTCCATTACACAGCTTTCTACGAGTCTGGTGGAGACTCCCTCATTAACGC 633
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 112 GluIlePheGlyGlyLeuAlaGlyPheLeuLeuLeuTrpSerAlaValLeuIleMetTy 131
QY 634 CTGAGCTACTGCTGTGATATCCCTGGCATTTGGAGCTACATCTGGAACCAATTTT 693
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 132 ProThrSerLeuAlaValIleSerMetThrPheSerAsnTyraValLeuGluProValPhe 151
QY 694 ATTCATGTGAATCCCTGAACCTTCGATCAAGCTCACTATTCAGCTGTGGGCATAACTGTA 753
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 152 ProAsnCysIleProProThrAlaSerArgValLeuSerMetAlaCysLeuMetLeu 171
QY 754 GTGATGTCCTAAATPAGCATGAGTGCAGTGGAGCCCGGATCCAGATTTTCTTAACC 813
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 172 LeuThrTrpValAsnSerSerValArgTrpAlaThrArgIleGluAspMetPheThr 191
QY 814 TTTTCGAAGCTCAGACAAATCTCATATTAATAGTCCCTGGAGTTATGACGCTAATAATAA 873
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 192 GlyGlyLysLeuAlaLeuSerLeuIleIleGlyValGlyLeuLeuGluIlePheGln 211
QY 874 GGTCAACGCGAGAACTTTTAA-----GAGCGCTTTTCA--GGAAGAGATCAAGTATT 924
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 212 GlyHisPheGluGluLeuArgProSerAsnAlaPheAlaPheTrpMetThrProSerVal 231
QY 925 ACGGGTTGGCCTGCTTTTATATGGAATGTATGCATATGCTGGCTGGTGTACCTC 984
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 232 GlyHisLeuAlaLeuAlaPheLeuGlnGlySerPheAlaPheSerGlyTrpAsnPheLeu 251
QY 985 AACTTTGTTACTGAAGAGTAGAAGAACCCCTGAGAAACCAATCCCTCCATATGATATA 1044
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
  
```

```

Db 252 AsnTyraValThrGluGluMetValAspAlaArgLysAsnLeuProArgAlaIlePheIle 271
QY 1045 TCCATGGCCATGTGCACCATTTGGCTATGCTGACAATGTGGCTACTTTAGACCAT 1104
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 272 SerIleProLeuValThrPheValTyThrPheThrAsnIleAlaTyThrAlaMet 291
QY 1105 AATGCTGAGGAGCTGCTGCTTTCAATGCAGTGGCAGTGCACCTTTCTGAGCGGCTACTG 1164
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 292 SerProGlnGluLeuLeuSerSerAsnAlaValAlaValThrPheGlyGluLysLeuLeu 311
QY 1165 GGAATTTCTCATTTAGCAGTTCGATCTTTGGCCCTCTCCCTGCTTGGCTCCATGAAC 1224
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 312 GlyTyraPheSerTrpValMetProValSerValAlaLeuSerThrPheGlyGlyIleAsn 331
QY 1225 GGTGCTGTGTTGCTCTCCAGTTCATGTTGTCGCTCTCGAGAGGGTCACTTCCA 1284
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 332 GlyTyraLeuPheThrTyraSerArgLeuCysPheSerGlyAlaArgGluGlyHisLeuPro 351
QY 1285 GAAATCTCTCCATGATTCATGTCGCAAGCACACACTCTCTACAGCTGTTATTTGTTG 1344
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 352 SerLeuLeuAlaMetIleHisValArgHisCysThrProIleProAlaLeuLeuValCys 371
QY 1345 CACCTTTTGACAATGATGATGCTCTCTGAGAGACTCGACAGCTTTTGAATTCCTC 1404
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 372 CysGlyAlaThrAlaValIleMetLeuValGlyAspThrTyraThrLeuIleAsnTyraVal 391
QY 1405 AGTTTTCAGGTGGCTTTTATTTGGCTGGCAGTGTGGCTGATTTATCTTCGATAC 1464
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 392 SerPheIleAsnTyraLeuCysTyraGlyValThrIleLeuGlyLeuLeuLeuArgTrp 411
QY 1465 AATGCCCAGATATGATGCTCTTTCAAGTGGCCACTGTTTCATCCAGCTTTGTTTCC 1524
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 412 ArgArgProAlaLeuHisArgProIleLysValAsnLeuLeuIleProValAlaTyraLeu 431
QY 1525 TTCACATGCTCTCTCATGCTTCCCTTTCCCTCTATTCGGACCCATTTAGTACAGGGAT 1584
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 432 ValPheTrpAlaPheLeuLeuValPheSerPheIleSerGluProMetValCysGlyVal 451
QY 1585 GGTTCGTCATCATCTGCTGAGTCCCTCGCTATTTATCTCTTTATTTATATGGACAAG 1644
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 452 GlyValIleIleIleLeuThrGlyValProIlePhePheLeuGlyValPheTrpArgSer 471
QY 1645 AAACCCAGGTGTTTAGAATAATGTCAGAGAAATAACCAACAACATTACAATAATACTG 1704
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 472 LysProLysCysValHisArgLeuThrGluSerMetThrHisTrpGlyGlnGluCys 491
QY 1705 GAAGTTGTA-----CCAGAAGAATAAGTTATGAACCTAATGGACTTGAG 1749
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 492 PheValValTyraProGlnAspAlaProGluGluGlu-----AsnGly----- 506
QY 1750 ATCTTTGGCAATCTGCCCA 1767
    |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:||| |||:|||
Db 507 -----ProCysPro 509
  
```

Search completed: April 16, 2003, 16:29:30
 Job time : 80.5 secs



Thu Apr 17 09:48:30 2003

cell culture; drug screening; gene therapy; antiinflammatory;
antisthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
antifungal; vulnery; antiulcer.

Homo sapiens.

W0200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001W0-US03800.

03-FEB-2000; 2000US-0496914.

27-APR-2000; 2000US-0560875.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-457740/49.

N-PSDB; ABA09201.

Human proteins and DNA encoding sequences useful for preventing,

treating or ameliorating a medical condition in a mammalian subject

e.g. arthritis and cancer --

Claim 20; Page 286; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activation; or may be involved in oncogenesis, cancer cell proliferation and metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a novel human polypeptide of the invention.

Sequence 521 AA;

Alignment Scores: 0 Length: 521
Pred. No.: 466.00 Matches: 466
Score: 100.00% Conservative: 0
Percent Similarity:

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	63.66%	Indels:	0
DB:	22	Gaps:	0
US-09-667-170a-440 (1-2239) x ABB11957 (1-521)			
QY	334	GAGAAAGTCAGCTGAAGAGAGAAAGTCACCTTTACTGAGGGAGTCTCCATTATCATTTGGC	393
DB	56	GLULYSVALGLNLEULYSARGLYSVALTHRLEULEUARGLYVALSERILEILEGLEY	75
QY	394	ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGCGGTCTCCAGAACACGGGAGC	453
DB	76	ThrIleilegleyAlaGlyIlePheIleSerProLYSGlyValLeuGlnAsnThrGlySer	95
QY	454	GTGGGCATGTCCTGACCATCTGGACGGTGTGTGGGGTCTCTCATCTATTGAGGCTTTG	513
DB	96	ValGlyMetSerLeuThrIleThrValCysGlyValLeuSerLeuPheGlyAlaLeu	115
QY	514	TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTACACATATATTTG	573
DB	116	SerTyrAlaGluLeuGlyThrThrIleLYSLysLysSerGlyIleHisTyrThrTyrIleLeu	135
QY	574	GAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCTCATATAATCCG	633
DB	136	GluValPheGlyProLeuProAlaPheValArgValTyrValGluLeuLeuIleArg	155
QY	634	CCTGCAGCTACTGCTGTATATCCCTGGCATTTGGAGCGGTACATTTCTGGAACCATTTT	693
DB	156	ProAlaAlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProPhePhe	175
QY	694	ATTCAATGTGAATCCCTGAACTTGGCATCAAGCTATTACAGCTGTGGGCATAATGTA	753
DB	176	IleGlnCysGluIleProGluLeuAlaIleLYSLysLeuIleThrAlaValGlyIleThrVal	195
QY	754	GTGATGTGCTCTAAATPAGCATGAGTGTACGTGAGCGCCCGGATCCAGATTCTTAACC	813
DB	196	ValMetValLeuAsnSerMetSerValSerTyrSerAlaArgIleGlnIlePheLeuThr	215
QY	814	TTTTGCAAGCTCACAGCAATCTGATAATATAGTCCCTGGAGTTATGCAGCTAATTA	873
DB	216	PheCysLYSLysLeuThrAlaIleLeuIleIleValProGlyValMetGlnLeuIleLYS	235
QY	874	GGTCAACACAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACCGGTG	933
DB	236	GlyGlnThrGlnAsnPhelysAspAlaPheSerGlyArgAspSerSerIleThrArgLeu	255
QY	934	CCACTGGCTTTTATATGGAATGATGATATGCTGCTGGCTGTTTACCTCAACTTTGTT	993
DB	256	ProLeuAlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTyrPheTyrLeuAsnPheVal	275
QY	994	ACTGAAGAAGTAGAAACCCCTGAAACACCATTCCTTGCATATGATATATCCATGCC	1053
DB	276	ThrGluGluValGluAsnProGluLYSThrIleProLeuAlaIleCysIleSerMetAla	295
QY	1054	ATTGTCACCATTTGGCTATGCTGACAAATGTGGCTACTTTTACGACCATTAATGCTGAG	1113
DB	296	IleValThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrThrIleAsnAlaGlu	315
QY	1114	GAGTGTGCTCTTCAATGCAAGTGGCAGTACCTTTCTGAGCGGCTACTGGGAATTC	1173
DB	316	GluLeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPhe	335
QY	1174	TCATTAGAGTTCGGATCTTTGTGTGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1233
DB	336	SerLeuAlaValProIlePheValAlaLeuSerCysPheLYSerMetAsnGlyIleVal	355
QY	1234	TTTGTGCTCTCAGGTTATCTATGTTTGGCTGCGAGGGGTCCACCTTCAGAAATCCTC	1293
DB	356	PheAlaValSerArgLeuPheTyrValAlaSerArgGlyHisLeuProGluIleLeu	375
QY	1294	TCCATGATTCATGTCGCGCAACACACCTCTTACACCTGTTATTTGTTTGCACCTTTG	1353
DB	376	SerMetIleHisValArgLYSHisThrProLeuProAlaValIleValLeuHisProLeu	395

XX The invention relates to isolated lung tumour-specific proteins and
CC their corresponding cDNA molecules. Lung tumour-specific proteins and
CC their antigen-presenting cells are useful for stimulating and/or
CC expanding T cells specific for a tumour protein, and for inhibiting
CC the development of cancer. The invention also relates to a composition
CC useful for stimulating an immune response, and for treating cancer. The
CC lung tumour specific oligonucleotide is useful in gene therapy and for
CC diagnosis, detection and treatment of lung cancer. The present sequence
CC is human T cell epitope related to lung tumour-specific protein.

XX Sequence 16 AA;
SQ

Alignment Scores:
Pred. No.: 16
Score: 16.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.19%
DB: 22

US-09-667-170A-440 (1-2239) x AAEL13851 (1-16)

QY 104 TTCCAGGCCAATTGTGGCATAGATTTTATCATATTCGTGATTTTGG 151
|||||
1 PheGlnAlaAsnCysGlyIleAspPheIlePheTrpIlePheTrp 16

Db

RESULT 4
AAV41159
ID AAY41159 standard; Protein; 511 AA.
XX
AC AAY41159;
XX
DT 24-JAN-2000 (first entry)
XX
DE Human amino acid permease homolog (AAPH).
XX
XX Amino acid permease homolog; AAPH: cancer; inflammatory disorder; AIDS;
KW autoimmune disorder; anemia; asthma; atherosclerosis; diabetes mellitus;
KW multiple sclerosis; osteoarthritis; osteoporosis; rheumatoid arthritis;
KW ulcerative colitis; infection; cell proliferation disorder; human.
XX
OS Homo sapiens.
XX
PN US981242-A.
XX
PD 09-NOV-1999.
XX
PF 03-FEB-1999; 99US-0243920.
XX
PR 06-MAY-1998; 98US-0073362.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Yue H, Corley NC, Hillman JL;
XX
XX WPI; 1999-633325/54.
DR N-PSDB; AAZ23240.
XX
XX New human amino acid permease homolog, useful in the diagnosis,
PT treatment and prevention of cancer, inflammatory/autoimmune disorders
PT and cell proliferation disorders -
PT
XX
XX Claim 1; Fig 1A-F; 32pp; English.
PS
XX
XX This represents a human amino acid permease homolog (AAPH). The AAPH
XX nucleic acid, protein, AAPH (ant)agonists are useful in the diagnosis,
CC treatment and prevention of cancer such as adenocarcinoma, leukemia, of the
CC lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and cancers of the
CC adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall
CC bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung,
CC muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands,
CC skin, spleen, testis, thymus, thyroid or uterus; inflammatory/autoimmune
CC

QY	1300	ATTCAATGCCCAAGCACACCTCTCTACCAAGCTGTTATTGTTTTCACCCCTTTTGACAATG	1359
Db	281	IleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeuThrMet	300
QY	1360	ATAATGCTCTCTCTCGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTTCGCCAGTGG	1419
Db	301	IleMetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPhLeuSerPheAlaArgTirp	320
QY	1420	CTTTTATTGGCGTGGCAGTGTCTGGCGCTGATTATCTTCGATACAAATGCCAGATATG	1479
Db	321	LeuPheIleGlyLeuAlaValAlaGlyLeuIleTyrLeuArgTyrLysCysProAspMet	340
QY	1480	CATCGTCTCTTCAAGTGCCACTGTTTCATCCCAGCTTTTCTTTCCTTCACATGCTCTTC	1539
Db	341	HisArgProPheLysValProLeuPheIleProAlaLeuPheSerPheThrCysLeuPhe	360
QY	1540	ATGGTTGGCTTTCCTCTATTCGGACCCATTTAGTACAGGATTGGCTTCCTCATCACT	1599
Db	361	MetValAlaLeuSerLeuTyrSerAspProPheSerThrGlyIleGlyPheValIleThr	380
QY	1600	CTGACTGGAGTCCCTCGCTATTATCTCTTATTATATGGACAAGAAACCCAGGTGGTTT	1659
Db	381	LeuThrGlyValProAlaTyrTyrLeuPheIleIleTirpAspLysLysProArgTirpPhe	400
QY	1660	AGAATAATCTCAGAGAAATACACAGACATTCACAAATACTGGAAGTTGTACCAGAA	1719
Db	401	ArgIleMetSerGluLysIleThrArgThrLeuGlnIleIleLeuGluValValProGlu	420
QY	1720	GAAGATAAGTTA	1731
Db	421	GluAspLysLeu	424
RESULT 3			
AAE13851	ID	AAE13851 standard; peptide: 16 AA.	
XX	AC	AAE13851;	
XX	DT	26-FEB-2002 (first entry)	
XX	DE	Human T cell epitope related to lung tumour-specific protein.	
XX	KW	Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;	
XX	KW	antisense-therapy; vaccine; immune response; lung cancer;	
XX	OS	T cell epitope.	
XX	PN	Homo sapiens.	
XX	PD	WO200172295-A2.	
XX	PF	04-OCT-2001.	
XX	PR	28-MAR-2001; 2001WO-US09991.	
XX	PR	29-MAR-2000; 2000US-0538037.	
XX	PR	05-JUN-2000; 2000US-0588957.	
XX	PR	18-AUG-2000; 2000US-0640878.	
XX	PR	22-SEP-2000; 2000US-234517P.	
XX	PR	01-NOV-2000; 2000US-0704512.	
XX	PR	14-DEC-2000; 2000US-0738973.	
XX	PA	(CORI-) CORIXA CORP.	
XX	PI	Reed SG, Lodes MJ, Mohamath R, Secretist H, Benson DR, Indirias CV;	
XX	PI	Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;	
XX	XX	WPI; 2001-639201/73.	
XX	DR		
XX	PT	New human lung-specific polynucleotides and polypeptides for the	
XX	PT	diagnosis and treatment of disease e.g. lung cancer	
XX	PS	Claim 2; Page 378; 378pp; English.	

disorders such as AIDS, Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, diabetes mellitus, Crohn's disease, atopic dermatitis, dermatomyositis, emphysema, multiple sclerosis, osteoarthritis, osteoporosis, rheumatoid arthritis, ulcerative colitis, viral, bacterial, fungal, parasitic, protozoal infections; and cell proliferation disorders such as actinic, keratosis, cirrhosis, arteriosclerosis, myelofibrosis, and primary thrombocythemia.

XX Sequence 511 AA;

Alignment Scores:
 Pred. No.: 0.00575 Length: 511
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 20 Gaps: 0

US-09-667-170A-440 (1-2239) x AAY41159 (1-511)

QY 517 TATGCTGAATGGGAACACTATAAGAAATCTGGA 552
 |||||
 Db 90 TyrAlaGluLeuGlyThrThrIleLysSerGly 101

RESULT 5

AAY27076
 ID AAY27076 standard; Protein: 511 AA.

XX AC AAY27076;

XX 18-OCT-1999 (first entry)

XX Human amino acid permease homologue (AAPH).

XX Amino acid permease homologue; AAPH; cancer; inflammatory disorder;
 KW autoimmune disorder; cell proliferation disorder; adenocarcinoma;
 KW AIDS; Crohn's disease; rheumatoid arthritis; arteriosclerosis; human;
 KW psoriasis; hepatitis; gene therapy.

XX Homo sapiens.

XX US5942399-A.

XX 24-AUG-1999.

XX 06-MAY-1998; 98US-0073362.

XX 06-MAY-1998; 98US-0073362.

XX (INCY-) INCYTE PHARM INC.

XX Corley NC, Hillman JL, Yue H;

XX WPI; 1999-493508/41.

XX N-PSDB; AAX86999.

XX New human amino acid permease homologue (AAPH) polynucleotide and its
 PT complement useful for prevention and treatment of cancer, autoimmune
 PT and cell proliferation disorders

XX Claim 1; Fig 1A-F; 33pp; English.

XX This sequence represents a human amino acid permease homologue (AAPH).
 CC The protein can be expressed by standard recombinant methodology. AAPH
 CC polynucleotides and polypeptides are useful for diagnosis, treatment and
 CC prevention of cancer e.g. adenocarcinoma, lung cancer, ovarian cancer;
 CC inflammatory/autoimmune disorders e.g. AIDS, Crohn's disease, rheumatoid
 CC arthritis; and cell proliferation disorders e.g. arteriosclerosis,
 CC psoriasis, hepatitis. AAPH polynucleotides are useful in gene therapy,
 CC in PCR-based methods for detecting upstream sequences, and also for

CC generating hybridisation probes for mapping the naturally occurring
 CC genomic sequence.

XX Sequence 511 AA;

Alignment Scores:
 Pred. No.: 0.00575 Length: 511
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 20 Gaps: 0

US-09-667-170A-440 (1-2239) x AAY27076 (1-511)

QY 517 TATGCTGAATGGGAACACTATAAGAAATCTGGA 552
 |||||
 Db 90 TyrAlaGluLeuGlyThrThrIleLysSerGly 101

RESULT 6

AAY87630

ID AAY87630 standard; Protein: 511 AA.

XX AC AAY87630;

XX 01-AUG-2000 (first entry)

XX Human monp-2 protein.

XX Human; monp-2; cell proliferation; drug development; immune disease;
 KW drug development; cancer.

XX Homo sapiens.

XX WO200017238-A1.

XX 30-MAR-2000.

XX 22-SEP-1999; 99WO-JP05205.

XX 24-SEP-1998; 98JP-0270356.

XX (TAIS) TAISHO PHARM CO LTD.

XX Yoshimoto M, Takayama K;

XX WPI; 2000-283547/24.

XX N-PSDB; AAX11954.

XX Novel gene monp-2 encoding monocyte-specific protein with cell
 PT proliferation promoting effect, useful in studying its expression and
 PT immune function, and in developing drugs for e.g. immune diseases and
 PT cancers

XX Claim 1a; Page 22-25; 34pp; Japanese.

XX This invention describes a novel protein MONP-2 (I) and its encoding
 CC gene monp-2 (II) with a cell proliferation promoting effect isolated
 CC from a human peripheral blood. The new protein (I) and gene are useful
 CC in studying the expression abnormality and dysfunction of (I) as well
 CC as maintenance of immune function and development of drugs for e.g.
 CC immune diseases and cancers. This represents encodes the MONP-2 protein
 CC described in the method of the invention.

XX Sequence 511 AA;

Alignment Scores:
 Pred. No.: 0.00575 Length: 511
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 21 Gaps: 0

US-09-667-170A-440 (1-2339) x AAY87630 (1-511)	
Qy	517 TATGTGTAATGGGAACAACATATAAAGAAATCTGGA 552
Ddb	90 TyrAlaGluLeuGlyThrIleLysLysSerGly 101
RESULT 7	
AAW93165	
ID	AAW93165 standard; Protein; 76 AA.
XX	
AC	AAW93165;
XX	
DT	05-NOV-2001 (first entry)
DE	Human digestive system antigen SEQ ID NO: 2514.
DE	
XX	
KK	Human; digestive system antigen; gene therapy; ulcerative colitis; infection; Hirschsprung's disease; digestive system disorder; Meckel's diverticulum
KW	
XX	
OS	Homo sapiens.
XX	
PN	WO200155314-A2.
XX	
PD	02-AUG-2001.
PF	
XX	17-JAN-2001; 2001WO-US01324.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205515.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0214886.
PR	30-JUN-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0216647.
PR	07-JUL-2000; 2000US-0216880.
PR	11-JUL-2000; 2000US-0217487.
PR	11-JUL-2000; 2000US-0217496.
PR	14-JUL-2000; 2000US-0218290.
PR	26-JUL-2000; 2000US-0220963.
PR	26-JUL-2000; 2000US-0220964.
PR	14-AUG-2000; 2000US-0224518.
PR	14-AUG-2000; 2000US-0224519.
PR	14-AUG-2000; 2000US-0225213.
PR	14-AUG-2000; 2000US-0225214.
PR	14-AUG-2000; 2000US-0225266.
PR	14-AUG-2000; 2000US-0225267.
PR	14-AUG-2000; 2000US-0225268.
PR	14-AUG-2000; 2000US-0225270.
PR	14-AUG-2000; 2000US-0225447.
PR	14-AUG-2000; 2000US-0225757.
PR	14-AUG-2000; 2000US-0225758.
PR	14-AUG-2000; 2000US-0225759.
PR	18-AUG-2000; 2000US-0226279.
PR	22-AUG-2000; 2000US-0226681.
PR	22-AUG-2000; 2000US-0226686.
PR	22-AUG-2000; 2000US-0227182.
PR	23-AUG-2000; 2000US-0227009.
PR	30-AUG-2000; 2000US-0228924.
PR	01-SEP-2000; 2000US-0229287.
PR	01-SEP-2000; 2000US-0229343.
PR	01-SEP-2000; 2000US-0229344.
PR	01-SEP-2000; 2000US-0229345.
PR	05-SEP-2000; 2000US-0229509.
PR	05-SEP-2000; 2000US-0229513.
PR	06-SEP-2000; 2000US-0230437.
PR	06-SEP-2000; 2000US-0230438.

XX The present sequence represents a sodium independent small neutral amino
 CC acid transporter polypeptide. The polypeptides and polynucleotides
 CC are used for controlling cell resistance to oxidizing stress, NMDA type
 CC glutamate receptor activity, cell growth and for screening for
 CC substances that have these activities.

XX Sequence 530 AA;

Alignment Scores: 530 AA;
 Pred. No.: 0.618 Length: 530
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.37% Indels: 0
 DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAG63803 (1-530)

QY 409 GGAATCTTCATCTCTCTAGGCGTGCTC 438

Db 63 GlyilePheileSerProlyGlyValLeu 72

RESULT 10

AAO12595

ID AAO12595 standard; Protein: 43 AA.

XX AAO12595;

XX 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 26487.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR N-PSDB; AA192526.

XX Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -

XX Claim 20; SEQ ID NO 26487; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 43 AA;

Alignment Scores:
 Pred. No.: 9.27 Length: 43
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.26% Indels: 0
 DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAO12595 (1-43)

QY 54 GACAGATCTCACTGTCACCCAGGCTG 28

Db 1 AspArgValSerLeuSerProArgLeu 9

RESULT 11

AAO09529

ID AAO09529 standard; Protein: 62 AA.

XX AAO09529;

XX 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 23421.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR N-PSDB; AAI89460.

XX Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -

XX Claim 20; SEQ ID NO 23421; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 62 AA; Alignment Scores: Pred. No.: 8.78 Length: 62 Score: 9.00 Matches: 9 Percent Similarity: 100.00% Conservative: 0 Best Local Similarity: 100.00% Mismatches: 0 Query Match: 1.26% Indels: 0 DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x AAO09529 (1-62)

QY 29 TGCATGATCTCTGCTCACTTCAACT 3
|||||
DB 7 TrpHisAspLeuCySerLeuGlnPro 15

RESULT 12
ABG19765
ID ABG19765 standard; Protein; 87 AA.
XX AC ABG19765;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #19756.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX N-PSDB; AAS83952.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 20; SEQ ID No 50124; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 87 AA; Alignment Scores: Pred. No.: 8.36 Length: 87 Score: 9.00 Matches: 9 Percent Similarity: 100.00% Conservative: 0 Best Local Similarity: 100.00% Mismatches: 0 Query Match: 1.23% Indels: 0 DB: 22 Gaps: 0

US-09-667-170A-440 (1-2239) x ABG19765 (1-87)

QY 72 AAAAAGAAAGAAAGAAAGAAAGAGAGA 98
|||||
DB 18 LysLysLysGluArgLysArgGluArg 26

RESULT 13
AAM85927
ID AAM85927 standard; Protein; 90 AA.
XX AC AAM85927;
XX DT 07-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen SEQ ID NO:13520.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 14-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.

Thu Apr 17 09:48:30 2003

PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0223287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-02321968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234957.
 PR 25-SEP-2000; 2000US-0234958.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236357.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-483426/52.
 DR N-PSDB; AAK58708.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Claim 11; SEQ ID NO 13520; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 90 AA;
 Alignment Scores:
 Pred. No.: 8 32 Length: 90
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.26% Indels: 0
 DB: 22 Gaps: 0

29 TGGCATGATCTCTGCTCACTTCAACCT 3

0
app.

29 TGGCATGATCTCTGCTCACTTCAACCT 3

DB:

US-09-667-170A-440 (1-2239) x AAY38429 (1-132)

us-09-667-170a-440.olin2p.rag

Thu Apr 17 09:48:30 2003

OY 1993 TTAGGAGAAAGACTAGACAATTACTA 2019
|||||
Db 34 LeuGlyGlulysThrArgGlnLeuLeu 42

Search completed: April 16, 2003, 16:35:36
Job time : 77.5 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:29:36 ; Search time 43.5 Seconds
(without alignments)
9896.320 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 732

Sequence: 1 ggagggtgaagtgcagag.....ttatttaaaaaaaaaaaaaa 2239

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 56536

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-O=/cgn2.1/USPTO_spool/US09667170/runat_04042003_091002_21212/app_query.fasta_1.2375
-DB=PIR.73 -OFT=FASTAN -SUFFIX=oln2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -NCPU=6 -ICPU=3
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -XGAPEXT=60 -XGAPEXT=60 -XGAPEXT=60
-XGAPOP=60 -XGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

PIR.73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	1.1	46	1	VCBPPF
2	8	1.1	67	2	A70375
3	8	1.1	82	2	S15143
4	8	1.1	82	2	C83556
5	8	1.1	100	2	H95283
6	8	1.1	128	2	S75737
7	8	1.1	159	2	T05279
8	8	1.1	184	2	H96718
9	8	1.1	260	1	S11984
10	8	1.1	287	2	H75098
11	8	1.1	348	2	E82933
12	8	1.1	353	2	A42700
13	8	1.1	359	2	I50704
14	8	1.1	361	2	T51078

c 15	8	1.1	366	2	T27257	hypothetical prote
c 16	8	1.1	399	2	T23470	hypothetical prote
c 17	8	1.1	404	2	F86445	unknown protein [i
c 18	8	1.1	419	1	S70617	homeotic protein g
c 19	8	1.1	464	2	T28818	hypothetical prote
c 20	8	1.1	483	2	T31138	hypothetical prote
c 21	8	1.1	485	2	T27147	hypothetical prote
c 22	8	1.1	486	2	G86320	F6A14.19 protein -
c 23	8	1.1	487	2	A71407	probable Ste20-lik
c 24	8	1.1	499	2	G86538	leucyl aminopeptid
c 25	8	1.1	499	2	G72083	leucine aminopepti
c 26	8	1.1	505	2	AC3043	aldehyde dehydroge
c 27	8	1.1	507	2	JG0165	LAT1 protein - hum
c 28	8	1.1	511	2	A98243	probable aldehyde
c 29	8	1.1	537	2	T27612	hypothetical prote
c 30	8	1.1	537	2	T04822	hypothetical prote
c 31	8	1.1	541	2	T31554	hypothetical prote
c 32	8	1.1	546	2	T37665	hypothetical prote
c 33	8	1.1	551	2	F69282	probable t-complex
c 34	8	1.1	554	2	A71729	medium-chain acyl-
c 35	8	1.1	562	2	T16854	DNA repair protein
c 36	8	1.1	582	2	T23624	hypothetical prote
c 37	8	1.1	617	2	T23623	hypothetical prote
c 38	8	1.1	662	2	T32821	hypothetical prote
c 39	8	1.1	679	2	S43129	penicillin-binding
c 40	8	1.1	725	2	A47168	cardiac morphogene
c 41	8	1.1	731	2	C85060	hypothetical prote
c 42	8	1.1	746	2	S31816	genome polyprotein
c 43	8	1.1	760	2	T06291	extensin homolog T
c 44	8	1.1	786	2	I49274	mammary gland fact
c 45	8	1.1	793	2	S54772	mammary gland fact

ALIGNMENTS

RESULT 1

VCBPPF

coat protein - phage Pfl

C:Species: phage Pfl

C>Date: 24-Sep-1981 #sequence_revision 24-Sep-1981 #text_change 31-Dec-1993

C:Accession: A04229

R:Nakashima, Y.; Wiseman, R.L.; Konigsberg, W.; Marvin, D.A.

A:Title: Primary structure and sidechain interactions of Pfl filamentous bacterial vi

A:Reference number: A04229; MUID:75082355; PMID:1110754

A:Accession: A04229

A:Molecule type: protein

A:Residues: 1-46 <NAK>

C:Comment: Pfl is a filamentous bacteriophage with single-stranded DNA.

C:Superfamily: filamentous phage coat protein B

Alignment Scores:

Pred. No.:	55	Length:	46
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	1	Gaps:	0

US-09-667-170A-440 (1-2239) x VCBPPF (1-46)

QY 1432 CTGGCAGTTGCTGGCTGATTAT 1455

|||||

Db 33 LeuAlaValAlaGlyLeuIleTyr 40

RESULT 2

A70375

ribosomal protein S21 - Aquifex aeolicus

C:Species: Aquifex aeolicus

C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 13-Aug-1999

C:Accession: A70375

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

us-09-667-170a-440.olin2p.rpr

Thu Apr 17 09:48:32 2003

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: A70375

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-67 <AQF>

A:Cross-references: GB:AE000711; NID:g2983401; PIDN:AAC06990.1; PID:g2983406; GB:AE00065

A:Experimental source: strain VF5

C:Genetics:

A:Gene: rpsu

C:Superfamily: Escherichia coli ribosomal protein S21

Alignment Scores:	52.9	Length:	67
Pred. No.:	8.00	Matches:	8
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x A70375 (1-67)

QY 77 GAAAGAAAGAAAAGAGAGAGAG 100

Db 43 GUAAGLysLysLysArgGluArg 50

RESULT 3

S15143 major coat protein precursor - phase pf1

N:Alternate names: gene 8 protein

C:Species: phase pf1

C:Date: 18-Feb-1994 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999

C:Accession: S15143; S20699

R:Hall, D.F.; Short, N.J.; Perham, R.N.; Petersen, G.B.

J. Mol. Biol. 218, 349-364, 1991

A:Title: DNA sequence of the filamentous bacteriophage pf1.

A:Reference number: S15140; MUID:91186399; PMID:2010913

A:Accession: S15143

A:Molecule type: DNA

A:Residues: 1-82 <JMG>

A:Cross-references: EMBL:X52107; NID:g14829; PIDN:CAA36331.1; PID:g14833

A:Experimental source: ATCC 25102-B1

C:Superfamily: filamentous phage coat protein B

F:1-36/Domain: signal sequence #status predicted <SIG>

F:37-82/Product: major coat protein #status predicted <MAT>

Alignment Scores:	51.8	Length:	82
Pred. No.:	8.00	Matches:	8
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x S15143 (1-82)

QY 1432 CTGGCAGTTGCTGGCTGATTAT 1455

Db 69 LeuAlaValAlaGlyLeuIleTyr 76

RESULT 4

C83556

coat protein B of bacteriophage (Pf1) PA0723 [imported] - Pseudomonas aeruginosa (strain

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: C83556

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Rody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: C83556

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-82 <STO>

A:Cross-references: GB:AE004507; GB:AE004091; NID:g9946596; PIDN:AAG04112.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: coaB; PA0723

C:Superfamily: filamentous phage coat protein B

Alignment Scores:	51.8	Length:	82
Pred. No.:	8.00	Matches:	8
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x C83556 (1-82)

QY 1432 CTGGCAGTTGCTGGCTGATTAT 1455

Db 69 LeuAlaValAlaGlyLeuIleTyr 76

RESULT 5

H95283

conserved hypothetical protein Sma0325 [imported] - Sinorhizobium meliloti (strain 10

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: H95283

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B

; Kalman, S.; Keating, D.H.; Palm, C.; Beck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K

proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: H95283

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-100 <KUP>

A:Cross-references: GB:AE006469; PIDN:AAK64834.1; PID:g14523247; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau

heubert, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

C:Genetics:

A:Gene: Sma0325

A:Genome: plasmid

Alignment Scores:	50.8	Length:	100
Pred. No.:	8.00	Matches:	8
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x H95283 (1-100)

QY 1075 CTGCAATGTGGCTACTTTACG 1098

Db 4 LeuThrAsnValAlaTyrPheThr 11

RESULT 6

S73737

hypothetical protein slr0921 - Synecocystis sp. (strain PCC 6803)

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S75737

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, N. A. Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* s.

A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S75737

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-128 <KAN>

A;Cross-references: EMBL:D64003; GB:AB001339; MID:g1001200; PIDN:BAA10472.1; PID:g100123

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C;Superfamily: *Synechocystis* hypothetical protein slr0921

Alignment Scores:

Pred. No.:	49.5	Length:	128
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.12%	Indels:	0
DB:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x S75737 (1-128)

QY 349 TCAGCTGCACCTTCTCTCTCCG 326

DB 89 SerAlaLeuSerLeuPheLeu 96

RESULT 7

T05279

transcription factor ATB2 - *Arabidopsis thaliana*

N;Alternate names: protein T4L20.170

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000

C;Accession: T05279

R;Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De

ewes, H.W.; Mayer, K.F.X.; Schueller, C.

submitted to the Protein Sequence Database, September 1998

A;Reference number: Z15406

A;Accession: T05279

A;Molecule type: DNA

A;Residues: 1-159 <BEV>

A;Cross-references: EMBL:AL023094

A;Experimental source: cultivar Columbia; BAC clone T4L20

C;Genetics:

A;Gene: ATB2

A;Map position: 4

A;Note: T4L20.170

C;Superfamily: tobacco DNA-binding protein tbz17; fos/jun DNA-binding domain homology

Alignment Scores:

Pred. No.:	48.4	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.12%	Indels:	0
DB:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x T05279 (1-159)

QY 1724 TCTTCTCTGTACACTTCAGT 1701

DB 4 SerSerGlyThrThrSer 11

RESULT 8

H96718

ACHVA22c, 50565-49239 [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: H96718

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzkeker, M.; Salzman, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talbotker, M.; Wu, D.; Yu, G.; Fraser, C.W.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H96718
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-184 <STO>
A;Cross-references: GB:AE005173; NID:96665546; PIDN:AAF22915.1; GSPDB:GN00141
C;Genetics:

A;Gene: T6C23.10

A;Map position: 1

Alignment Scores:

Pred. No.:	47.7	Length:	184
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x H96718 (1-184)

QY 963 ATATGCTGGCTGTTTACCTCAA 986

DB 85 IllecysirpLeuValLeuProGln 92

RESULT 9

S11984

fapR protein - *Escherichia coli*

C;Species: *Escherichia coli*

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: S11984

R;Klaassen, P.; de Graaf, F.K.

Mol. Microbiol. 4, 1779-1783, 1990

A;Title: Characterization of fapR, a positive regulator of expression of the 987P operon

A;Reference number: S11984; MUID:91171879; PMID:2077360

A;Accession: S11984

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-260 <KLA>

A;Cross-references: EMBL:X53494

A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in lacking 96-

C;Genetics:

A;Gene: fapR

C;Superfamily: fapR protein

C;Keywords: DNA binding; transcription regulation

Alignment Scores:

Pred. No.:	46	Length:	260
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	1	Gaps:	0

US-09-667-170A-440 (1-2239) x S11984 (1-260)

QY 1135 GTGCAGTCACCTTTCTGAGCGG 1158

DB 148 ValAlaValThrPheSerGluArg 155

RESULT 10

H75098

hypothetical protein PAB0790 - *Pyrococcus abyssi* (strain Orsay)

C;Species: *Pyrococcus abyssi*

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

us-09-667-170a-440.olin2p.rpr

Thu Apr 17 09:48:32 2003

C:Accession: H75098
 R:Anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
 A:Reference number: A75001
 A:Accession: H75098
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-287 <KAW>
 A:Cross-references: GB:AJ248286; GB:AL096836; NID:g5459366; PIDN:CAB50093.1; PID:e151599
 A:Experimental source: strain Orsay
 C:Genetics:

Alignment Scores:
 Pred. No.: 45.6 Length: 287
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservatively: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: Gaps: 0

US-09-667-170A-440 (1-2239) x H75098 (1-287)

QY 1369 TTCTGTGGACCTCGACAGTCTT 1392

Db 53 PheSerGlyAspLeuAspSerLeu 60

RESULT 11

E82933
 Type I restriction enzyme M protein, truncated homolog U0098 [imported] - Ureaplasma urealyticum
 C:Species: Ureaplasma urealyticum
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 16-Aug-2002
 C:Accession: E82933
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to GenBank, February 2000
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
 A:Reference number: A82870
 A:Accession: E82933
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-348 <GLA>
 A:Cross-references: GB:AE002110; GB:AF222894; NID:g6899051; PIDN:AAF30504.1; GSPDB:GN001
 A:Experimental source: serovar 3; biovar 1
 C:Genetics:
 A:Gene: hsdM-1; U0098
 A:Genetic code: SGC3
 C:Superfamily: type I site-specific deoxyribonuclease chain hsdM (associate member)

Alignment Scores:
 Pred. No.: 44.7 Length: 348
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservatively: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.12% Indels: 0
 DB: Gaps: 0

US-09-667-170A-440 (1-2239) x E82933 (1-348)

QY 1733 CATACTATCTTCTTCTGTGACA 1710

Db 173 HisAsnLeuSerSerGlyThr 180

RESULT 12

A42700
 replication factor C - human
 C:Species: Homo sapiens (man)
 C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Aug-1998
 A:Accession: A42700
 R:Chen, M.; Pan, Z.Q.; Hurwitz, J.
 Proc. Natl. Acad. Sci. U.S.A. 89, 2516-2520, 1992
 A:Title: Sequence and expression in Escherichia coli of the 40-kDa subunit of activator
 A:Reference number: A42700; MUID:92212860; PMID:1313560

A:Accession: A42700
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-353 <CHE>
 A:Cross-references: GB:M87338
 C:Genetics:
 A:Gene: GDB:RFC2; A1
 A:Cross-references: GDB:283460
 A:Map position: 7q11.23-7q11.23
 C:Superfamily: phage T4 DNA polymerase accessory protein 44

Alignment Scores:
 Pred. No.: 44.6 Length: 353
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservatively: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: Gaps: 0

US-09-667-170A-440 (1-2239) x A42700 (1-353)

QY 1761 CTGCCAAGGGGAGACACAAATA 1784

Db 129 LeuProLysGlyArgHisLysIle 136

RESULT 13

I50704
 replication factor C/activator 1 subunit - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 20-Sep-1999
 C:Accession: I50704
 R:Oberholzer, J.; Cohen, E.L.; Davis, J.G.
 DNA Cell Biol. 13, 857-863, 1994
 A:Title: Molecular cloning of a chick cochlea cDNA encoding a subunit of DNA replicat
 A:Reference number: I50704; MUID:94347215; PMID:8068208
 A:Accession: I50704
 A:Status: preliminary; translated from GB/EMBL/DDJ
 A:Molecule type: mRNA
 A:Residues: 1-359 <OBE>
 A:Cross-references: EMBL:U12438; NID:g527668; PIDN:AAA20552.1; PID:g527669
 C:Superfamily: phage T4 DNA polymerase accessory protein 44

Alignment Scores:
 Pred. No.: 44.5 Length: 359
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservatively: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: Gaps: 0

US-09-667-170A-440 (1-2239) x I50704 (1-359)

QY 1761 CTGCCAAGGGGAGACACAAATA 1784

Db 135 LeuProLysGlyArgHisLysIle 142

RESULT 14

T51078
 related to homeotic protein Antennapedia [imported] - Neurospora crassa
 N:Alternate names: protein B2A19.120
 C:Species: Neurospora crassa
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T51078
 R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
 submitted to the Protein Sequence Database, July 2000
 A:Reference number: Z25286
 A:Accession: T51078
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-361 <SCH>
 A:Cross-references: EMBL:AL390092; GSPDB:GN00116; NCSP:B2A19.120
 A:Experimental source: BAC clone B2A19; strain OR74A
 C:Genetics:

A:Gene: NCSP:B2A19.120
A:Map position: 6
A:Introns: 282/2

Alignment Scores:

Pred. No.:	44.5	Length:	361
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x T51078 (1-361)

QY 148 TTGGATTCTTTGTTTCTCATCA 171

Db 311 LeuAspSerPheValPheSerSer 318

RESULT 15

T27257

hypothetical protein Y62H9A.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T27257

R:Steward, C.

submitted to the EMBL Data Library, October 1998

A:Reference number: 220334

A:Accession: T27257

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-366 <WIL>

A:Cross-references: EMBL:AL032630; PIDN:CAA21558.1; GSPDB:GN00028; CESP:Y62H9A.1

A:Experimental source: clone Y62H9A

C:Genetics:

A:Gene: CESP:Y62H9A.1

A:Map position: X

A:Introns: 25/3; 206/3; 287/2; 342/1

Alignment Scores:

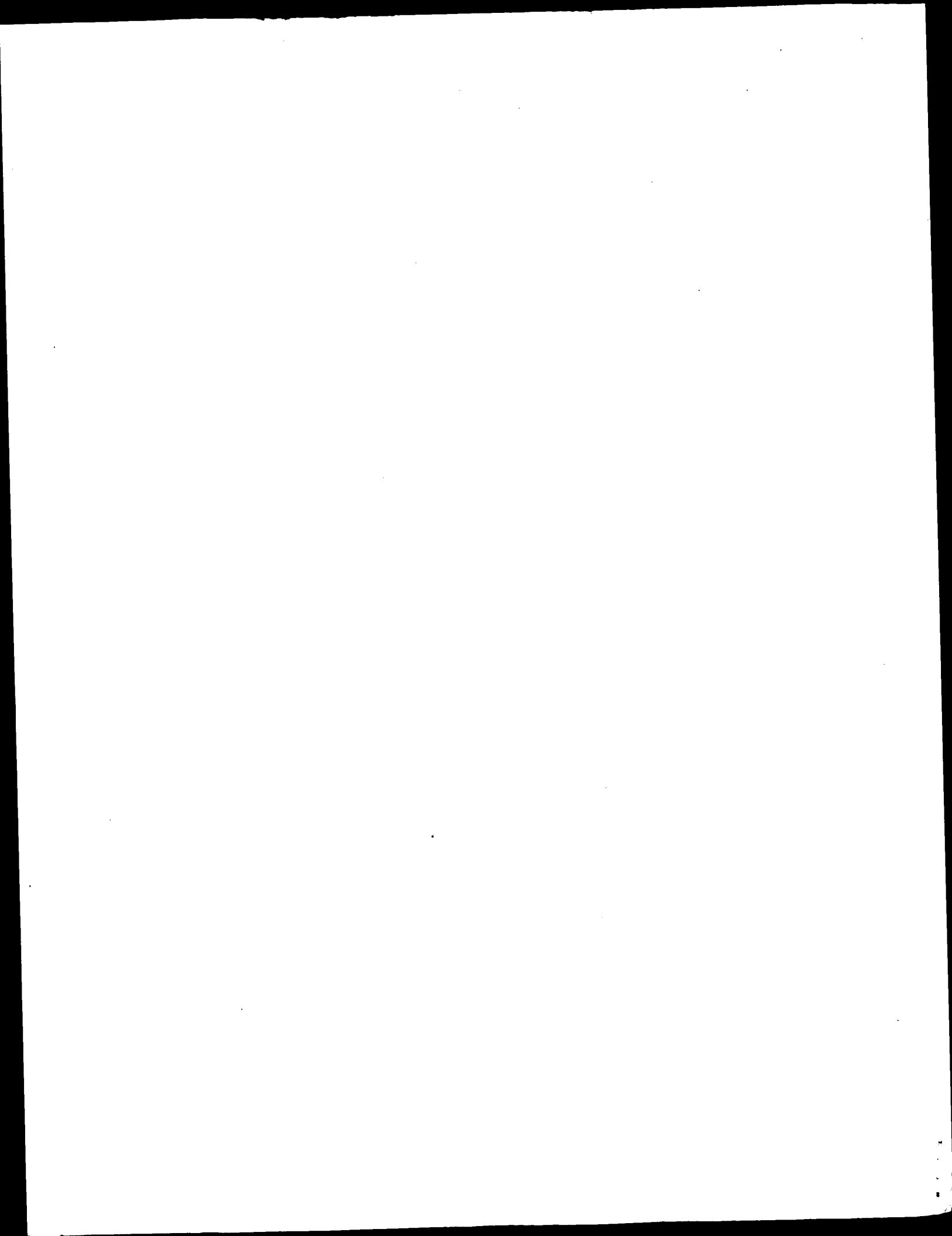
Pred. No.:	44.4	Length:	366
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-09-667-170A-440 (1-2239) x T27257 (1-366)

QY 2082 TTTTCTGTATATGGTTTCTTA 2105

Db 71 PheSerValrMetGlyPheVal 78

Search completed: April 16, 2003, 16:41:50
Job time : 48.5 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:26:10 ; Search time 25.5 Seconds
(without alignments)
7283.569 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 732

Sequence: 1 ggagggtgaagtgcagcag.....ttatttaaaaaaaaaaaaaa 2239

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+n2p.model -DEV-xmlh
-Q/cgn2_1/USPTO_spool/US09667170/runat_04042003_091001_21192/app_query.fasta_1.2375
-DB-SwissProt_40 -Qfmt-fastan -SUFFIX-olin2p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX-oligo -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE-quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09667170@cgn_1_1_24_@runat_04042003_091001_21192 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGESQUEY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -XGAPEXT=60 -XGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Swissprot_40:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	466	63.7	501	1 XCT_HUMAN	Q9upv5 homo sapien
2	46	6.3	502	1 XCT_MOUSE	Q9wtr6 mus musculus
3	12	1.6	511	1 YLAI_HUMAN	Q9umr0 homo sapien
4	8	1.1	67	1 RS21_AQUAE	Q67028 aquifex aeo
5	8	1.1	82	1 COAB_BPPFI	P03621 bacterioph
6	8	1.1	260	1 FAPR_ECOLI	P23774 escherichia
7	8	1.1	354	1 AC14_HUMAN	P35250 homo sapien
8	8	1.1	419	1 GSC_DROME	P54366 drosophila
9	8	1.1	467	1 2A5E_HUMAN	Q16537 h serine/th
10	8	1.1	499	1 AMPA_CHLPN	Q928f8 chlamydia p
11	8	1.1	507	1 LATI_HUMAN	Q01650 homo sapien
12	8	1.1	512	1 LATI_MOUSE	Q92127 mus musculus
13	8	1.1	512	1 LATI_RAT	Q63016 rattus norv
14	8	1.1	512	1 LATI_MOUSE	Q9qxw9 mus musculus
15	8	1.1	531	1 LAT2_MOUSE	Q9wtr6 rattus norv
16	8	1.1	531	1 LAT2_RAT	Q9uh15 homo sapien
17	8	1.1	535	1 LAT2_HUMAN	Q9utm4 schizosacch
18	8	1.1	546	1 TCPE_SCHPO	

19	8	1.1	554	1	RECN_RICPR	Q9zdq2 rickettsia
20	8	1.1	786	1	ST5B_MOUSE	P42232 mus musculus
21	8	1.1	786	1	ST5B_RAT	P52632 rattus norv
22	8	1.1	793	1	ST5A_MOUSE	P42230 mus musculus
23	8	1.1	793	1	ST5A_RAT	O62771 rattus norv
24	8	1.1	1447	1	SGS1_YEAST	P35187 saccharomyc
25	8	1.1	1483	1	UFDA_YEAST	P33202 saccharomyc
26	8	1.1	2226	1	POLG_HPAV2	P26580 hepatitis a
27	8	1.1	2226	1	POLG_HPAV4	P26581 hepatitis a
28	8	1.1	2227	1	POLG_HPAV8	P26582 hepatitis a
29	8	1.1	2227	1	POLG_HPAVH	P06417 hepatitis a
30	8	1.1	2227	1	POLG_HPAVL	P06441 hepatitis a
31	7	1.0	45	1	Y843_HAEIN	P44059 haemophilus
32	7	1.0	56	1	Y048_BORBU	O51077 borrelia bu
33	7	1.0	104	1	C552_BRAJA	P24059 bradyrhizob
34	7	1.0	108	1	CYC_CURLU	Q96vp3 curvularia
35	7	1.0	112	1	CYC_ARATH	P29380 arabidopsis
36	7	1.0	127	1	SECE_ECOLI	P16920 escherichia
37	7	1.0	127	1	SECE_SALTY	O919k1 salmonella
38	7	1.0	129	1	RS9_CHLTR	O84128 chlamydia t
39	7	1.0	130	1	INT2_HUMAN	Q28808 pan troglod
40	7	1.0	130	1	INT2_PANTR	Q9pkr2 chlamydia m
41	7	1.0	133	1	RS9_CHLMU	Q08535 mus musculus
42	7	1.0	133	1	SECR_MOUSE	Q928t8 chlamydia p
43	7	1.0	134	1	RS9_CHLPN	P26533 synecocyst
44	7	1.0	135	1	ATPE_SYNY3	P03167 woodchuck h
45	7	1.0	141	1	X_WHV1	

ALIGNMENTS

RESULT 1	XCT_HUMAN	STANDARD	PRT	501 AA
ID	Q9UPV5			
AC	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Cystine/glutamate transporter (Amino acid transport system xc-) (XCT)			
DE	(Calcium channel blocker resistance protein CCBRI).			
GN	SLC7A11.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21082210; PubMed=11213471;			
RA	Sato H., Tamba M., Kuriyama-Matsumura K., Okuno S., Bannai S.;			
RT	"Molecular cloning and expression of human XCT, the light chain of			
RT	amino acid transport system xc-";			
RL	Antioxid. Redox. Signal. 2:665-671(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Conklin D.S., Beach D.H.;			
RT	"CCBR1, novel CD98 light chain implicated in redox control and calcium			
RT	signaling.";			
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RA	Chaney C.D., Kekuda R., Wang H., Huang W., Prasad P.D., Smith S.B.,			
RA	Ganapathy V.;			
RT	"Structure, function and regulation of human cystine/glutamate			
RT	transporter in retinal pigment epithelial cells.";			
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Borsani G., Manzoni M., Palacin M., Pineda M., Gasol E.;			
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			

RA Strausberg R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY EXCHANGE OF ANIONIC
 CC AMINO ACIDS WITH HIGH SPECIFICITY FOR ANIONIC FORM OF CYSTINE AND
 CC GLUTAMATE (BY SIMILARITY).
 CC - SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
 CC TRANSPORT PROTEIN SLC3A2/4F2HC (BY SIMILARITY).
 CC - SUBCELLULAR LOCATION: Integral membrane protein (probable).
 CC - SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
 CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB026891; BAA82628.1; -
 DR EMBL; AF200708; AAG35592.1; -
 DR EMBL; AF252872; AAK49111.1; -
 DR EMBL; AJ277882; CAC81905.1; -
 DR EMBL; BC012087; AAK12087.1; -
 DR EMBL; HGNC:11059; SLC7A11.
 DR InterPro; IPR002293; AA/re1_primease1.
 DR InterPro; IPR004760; L_AA_transport.
 DR InterPro; IPR004841; Permease.
 DR Pfam; PF00324; aa_permeases; 1.
 DR TIGRFAMs; TIGR00911; 2A0308; 1.
 DR Transport; Amino-acid transport; Transmembrane.
 KW TRANSMEM 44 64
 FT TRANSMEM 76 96
 FT TRANSMEM 136 156
 FT TRANSMEM 159 179
 FT TRANSMEM 190 210
 FT TRANSMEM 235 255
 FT TRANSMEM 266 286
 FT TRANSMEM 318 338
 FT TRANSMEM 365 385
 FT TRANSMEM 388 408
 FT TRANSMEM 423 443
 FT TRANSMEM 450 470
 FT TRANSMEM 501 521
 FT CARBOHYD 314 314
 SQ SEQUENCE 501 AA; 55422 MW; 3EF2648B94A9F59B CRC64;

Alignment Scores:
 Pred. No.: 0 Length: 501
 Score: 466.00 Matches: 466
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 63.66% Indels: 0
 DB: 1 Gaps: 0

US-09-667-170a-440 (1-2239) x XCT_HUMAN (1-501)
 QY 334 GAGAAAGTCAGCTGAAGAGAAAGTCATCTTACTCAGGGAGTCTCCATTATCATTTGGC 393
 DB 36 GlulysValGlnLeulysArgLysValThrLeuLeuArgGlyValSerIleIleGly 55
 QY 394 ACCATATTCGAGCAGAAATCTTCATCTCTCCTAAGGGGCTCTCAGAACACGGGACG 453
 DB 56 ThrIleIleGlyAlaGlyIlePheIleSerProLysGlyValLeuGlnAsnThrGlySer 75
 QY 454 GTGGGCTGTCCTGACCATCTCGAGCGGTGTGTGGGGTCTCTGCTACTATTTGGAGCTTTG 513
 DB 76 ValGlyMetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuPheGlyAlaLeu 95
 QY 514 TCTTATGCTGAATGGGAACTATTAAGAAATCTGGAGGTCATTACACATATTTTGG 573
 DB 96 SerTyrAlaGluLeuGlyThrThrIleLysLysSerGlyCysIleThrThrIleLeu 115

QY 574 GAAGTCCTTTGGTCCATTACACAGCTTTTCTAGAGTCTGGGTGGAACTCCTCATATACGC 633
 DB 116 GluValPheGlyProLeuProAlaPheValArgValTrpValGluLeuLeuIleArg 135
 QY 634 CCTGACGCTACTGCTGTGATATCCCTGGCATTGGACCTACATCTGGAAACCACTTTT 693
 DB 136 ProAlaAlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProPhePhe 155
 QY 694 ATTCATATGTAATCCTGAATCTGCATCAAGCTCAAGCTCATTCAGCTGTGGGCACTACTGTA 753
 DB 156 IleGlnCysGluIleProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrVal 175
 QY 754 GTGATGCTCTAAATAGATGATGTCAGCTGGAGCGCCGATCCAGATTTCTTAACC 813
 DB 176 ValMetValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThr 195
 QY 814 TTTTGAAGCTCAGCAATCTCGATAATATAGTCCCTGGAGTTATGACGCTAATATAA 873
 DB 196 PheCysLysLeuThrAlaIleLeuIleIleValProGlyValMetGlnLeuIleLys 215
 QY 874 GGTCAAACGAGAACTTTAAAGACGCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTG 933
 DB 216 GlyGlnThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerSerIleThrArgLeu 235
 QY 934 CCAGTGGCTTTTATTTATGGAATGTATGATGCTGGCTGCTTTTACCTCAACTTTTGT 993
 DB 236 ProLeuAlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTrpPheTyrLeuAsnPheVal 255
 QY 994 ACTGAAGAAGTAGAAAACCTGAAAACCAATCCCTCTGCAATATGATATATCCATGGCC 1053
 DB 256 ThrGluGluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAla 275
 QY 1054 ATTCTCAGCTATGGCTATGTGTCAGCAAAATGCTGCTACTTTACGACCACTTATGCTGAG 1113
 DB 276 IleValThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrIleAsnAlaGlu 295
 QY 1114 GAGTGTGCTTTTCAAATGAGTGGCAGTACCTTTTCTGAGCGGTACTGGGAAATTC 1173
 DB 296 GluLeuLeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPhe 315
 QY 1174 TCATTAGCAGTTCGAGTCTTTGTTGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1233
 DB 316 SerLeuAlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyGlyVal 335
 QY 1234 TTTGCTGCTCCAGTATTCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1293
 DB 336 PheAlaValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIleLeu 355
 QY 1294 TCCATGATTCATGTCGCAAGCACACTCTCTACAGCTGTTATTGTTTGGACCTTTG 1353
 DB 356 SerMetIleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeu 375
 QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTGACAGTCTTTTGAATTTCTCAGTTTGGC 1413
 DB 376 ThrMetIleMetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPheLeuSerPheAla 395
 QY 1414 AGTGGCTTTTATTGGGCTGGCAGTCTGCTGGCTGATTATCTTCATACAAATGCCA 1473
 DB 396 ArgTrpLeuPheIleGlyLeuAlaValAlaGlyLeuIleTyrLeuArgTyrLysCysPro 415
 QY 1474 GATATGATTCCTCTTCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1533
 DB 416 AspMetHisArgProPheLysValProLeuPheIleProAlaLeuPheSerPheThrCys 435
 QY 1534 CTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593
 DB 436 LeuPheMetValAlaLeuSerLeuTyrSerAspProPheSerThrGlyIleGlyPheVal 455
 QY 1594 ATCACTCTGCTGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1653
 DB 456 IleThrLeuThrGlyValProAlaTyrTyrLeuPheIleIleTrpAspLysLysProArg 475
 QY 1654 TGGTTTGAATAATGTGACAGAAAATAACAGAACATTACAATAATTAATGGAAGTTGTA 1713

Db 476 TtpPheArgIleMetSerGluLysIleThrArgThrLeuGlnIleLeuGluVal 495
 QY 1714 CCAGAGAGATAGTTA 1731
 Db 496 ProGluGluAspLysLeu 501
 RESULT 2
 XCT_MOUSE
 ID XCT_MOUSE STANDARD; PRT; 502 AA.
 AC Q9WTR6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Cystine/glutamate transporter (Amino acid transport system xc-) (XCT).
 GN SLC7A11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. PubMed=10206947;
 RX MEDLINE=99223452; PubMed=10206947;
 RA Sato H., Tamba M., Ishii T., Bannai S.;
 RT "Cloning and expression of a plasma membrane cystine/glutamate
 exchange transporter composed of two distinct proteins.";
 RL J. Biol. Chem. 274:11455-11458(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ; TISSUE=Liver;
 RA Sasaki H., Sato H., Bannai S.;
 RT "Isolation and functional characterization of mouse cystine/glutamate
 exchange transporter gene.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY EXCHANGE OF ANIONIC
 CC AMINO ACIDS WITH HIGH SPECIFICITY FOR ANIONIC FORM OF CYSTINE AND
 CC GLUTAMATE.
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
 CC TRANSPORT PROTEIN SLC3A2/4F2HC.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
 CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
 CC SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AB022345; BAA77220.1;
 DR EMBL; AB037661; BAA05522.1;
 DR EMBL; AB037650; BAA05522.1; JOINED.
 DR EMBL; AB037651; BAA05522.1; JOINED.
 DR EMBL; AB037652; BAA05522.1; JOINED.
 DR EMBL; AB037653; BAA05522.1; JOINED.
 DR EMBL; AB037654; BAA05522.1; JOINED.
 DR EMBL; AB037655; BAA05522.1; JOINED.
 DR EMBL; AB037656; BAA05522.1; JOINED.
 DR EMBL; AB037657; BAA05522.1; JOINED.
 DR EMBL; AB037658; BAA05522.1; JOINED.
 DR EMBL; AB037659; BAA05522.1; JOINED.
 DR EMBL; AB037660; BAA05522.1; JOINED.
 DR MGD; MGI:1347355; SLC7A11.
 DR InterPro; IPR002293; AA/rel_pmease1.
 DR InterPro; IPR004760; LAA_transport.
 DR InterPro; IPR004841; Permease.
 DR Pfam; PF00324; aa_permeases; 1.
 DR TIGRFAMs; TIGR00911; 2A0308; 1.
 KW Transport; Amino-acid transport; Transmembrane.
 FT TRANSMEM 44 64 POTENTIAL.

FT TRANSMEM 76 96 POTENTIAL.
 FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 159 179 POTENTIAL.
 FT TRANSMEM 190 210 POTENTIAL.
 FT TRANSMEM 235 255 POTENTIAL.
 FT TRANSMEM 266 286 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT TRANSMEM 365 385 POTENTIAL.
 FT TRANSMEM 388 408 POTENTIAL.
 FT TRANSMEM 423 443 POTENTIAL.
 FT TRANSMEM 450 470 POTENTIAL.
 FT CARBOHYD 19 19 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 502 AA; 55455 MW; AA185102D83992A CRC64;
 Alignment Scores:
 Pred. No.: 2.02e-39 Length: 502
 Score: 46.00 Matches: 46
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.28% Indels: 0
 DB: 1 Gaps: 0
 US-09-667-170A-440 (1-2239) x XCT_MOUSE (1-502)
 QY 1171 TTCTCATTAGCAGTTCGGATCTTTGTTCCTCTCTGCTTTCCTCCATGACGGTGGT 1230
 Db 315 PheserLeuAlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyGly 334
 QY 1231 GTGTTTCTGCTCTCCAGGTTTATTCTATGTTGCGTCTCGAGAGGGTCACCTTCCAGAAATC 1290
 Db 335 ValPheAlaValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIle 354
 QY 1291 CTCTCCATGATTTCATGTC 1308
 Db 355 LeuSerMetIleHisVal 360
 RESULT 3
 YLAI_HUMAN
 ID YLAI_HUMAN STANDARD; PRT; 511 AA.
 AC Q9UM01; O9P2V5; O95984;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Y+L amino acid transporter 1 (Y(+)-L-type amino acid transporter 1)
 DE (Y+LAT-1) (Y+LAT1) (Monocyte amino acid permease 2) (MOP-2).
 GN SLC7A7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT LPI ARG-334.
 RX MEDLINE=99047611; PubMed=9829974;
 RA Torrens D., Estevez R., Pineda M., Fernandez E., Lloberas J.,
 RA Shi Y.-B., Zorzano A., Palacin M.;
 RT "Identification and characterization of a membrane protein (Y+L amino
 RT acid transporter-1) that associates with 4F2hc to encode the amino
 RT acid transport activity Y+L. A candidate gene for lysinuric protein
 RT intolerance.";
 RL J. Biol. Chem. 273:32437-32445(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=9904891; PubMed=9878049;
 RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;
 RT "Amino acid transport of Y+L-type by heterodimers of 4F2hc/CD98 and
 RT members of the glycoprotein-associated amino acid transporter
 RT family.";
 RL EMBO J. 18:49-57(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;

RX MEDLINE=99178268; PubMed=10080183;
RA Borsani G., Bassi M.T., Sperandio M.P., De Grandi A., Buoninconti A.,
RA Riboni M., Manzoni M., Incerti B., Pepe A., Andria G., Ballabio A.,
RA Sebastio G.;
RT "SLC7A7, encoding a putative permease-related protein, is mutated in
RL patients with lysinuric protein intolerance.";
RL Nat. Genet. 21:297-301(1999).
RN [4]
RN SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RC Fukasawa Y., Segawa H., Endou H., Kanai Y.;
RA "Characterization of a human system y+L amino acid transporters.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RX MEDLINE=20202143; PubMed=10737982;
RA Noguchi A., Shoji Y., Koizumi A., Takahashi T., Shoji Y.,
RA Matsumori M., Kayo T., Onata T., Wada Y., Yoshimura I., Maisawa S.,
RA Konishi M., Takasago Y., Takada G.;
RT "SLC7A7 genomic structure and novel variants in three Japanese
RL lysinuric protein intolerance families.";
RL Hum. Mutat. 15:367-372(2000).
RN [6]
RN SEQUENCE FROM N.A.
RA Takayama K., Yoshimoto M.;
RT "Molecular and biological characterization of a novel monocyte amino
RL acid permease, MOP-2.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [8]
RN MEDLINE=20100611; PubMed=106311139;
RA Sperandio M.P., Bassi M.T., Riboni M., Parenti G., Buoninconti A.,
RA Manzoni M., Incerti B., Larocca M.R., Di Rocco M., Strisciuglio P.,
RA Dianzani U., Parini R., Candito M., Endo F., Ballabio A., Andria G.,
RA Sebastio G., Borsani G.;
RT "Structure of the SLC7A7 gene and mutational analysis of patients
RL affected by lysinuric protein intolerance.";
RL Am. J. Hum. Genet. 66:92-99(2000).
RN [9]
RN VARIANTS LPI VAL-54 AND ASP-338.
RX MEDLINE=20122253; PubMed=10655553;
RA Mykkaenen J., Torrens D., Pineda M., Camps M., Yoldi M.E., Simell O.,
RA Horelli-Kuitunen N., Huoponen K., Heinoonen M., Oksanen J., Simell O.,
RA Savontaus M.-L., Zorzano A., Palacin M., Aula P.;
RT "Functional analysis of novel mutations in y+LAT-1 amino acid
RL transporter gene causing lysinuric protein intolerance (LPI).";
RL Hum. Mol. Genet. 9:431-438(2000).
CC -!- FUNCTION: SODIUM-INDEPENDENT EXCHANGER OF CATIONIC AND LARGE
CC NEUTRAL AMINO ACIDS.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC TRANSPORTER PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY >> PERIPHERAL BLOOD
CC LEUKOCYTES >> LUNG > PLACENTA = SPLEEN > SMALL INTESTINE.
CC -!- DISEASE: DEFECTS IN SLC7A7 ARE A CAUSE OF LYSINURIC PROTEIN
CC INTOLERANCE (LPI). LPI IS AN AUTOSOMAL RECESSIVE MULTISYSTEM
CC DISORDER FOUND MAINLY IN FINLAND AND ITALY. ON A NORMAL DIET, LPI
CC PATIENTS PRESENT POOR FEEDING, VOMITING, DIARRHEA, EPISODES OF
CC HYPERAMMONAEMIC COMA AND GROWTH RETARDATION. HEPATOSPLENOMEGALY,
CC OSTEOPOROSIS AND A LIFE-THREATENING PULMONARY INVOLVEMENT
CC (ALVEOLAR PROTEINOSIS) ARE ALSO SEEN. BIOCHEMICALLY LPI IS
CC CHARACTERIZED BY A DEFECT IN THE PLASMA MEMBRANE TRANSPORT OF
CC DIBASIC AMINO ACIDS.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC SUBFAMILY.
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y18474; CAB40136.1; -
DR EMBL; AJ130718; CAA10198.1; -
DR EMBL; AF092032; AAC83706.1; -
DR EMBL; AB020532; BAA87623.1; -
DR EMBL; AB031537; BAA95120.1; ALT_SEQ.
DR EMBL; AB031529; BAA95120.1; JOINED.
DR EMBL; BAA31530; BAA95120.1; JOINED.
DR EMBL; AB031531; BAA95120.1; JOINED.
DR EMBL; AB031532; BAA95120.1; JOINED.
DR EMBL; AB031533; BAA95120.1; JOINED.
DR EMBL; AB031534; BAA95120.1; JOINED.
DR EMBL; AB031535; BAA95120.1; JOINED.
DR EMBL; AB031536; BAA95120.1; JOINED.
DR EMBL; AB011263; BAB11849.1; -
DR EMBL; BC003062; AAH03062.1; -
DR EMBL; BC010107; AAH10107.1; -
DR Gene; HGNC:11065; SLC7A7.
DR MIM; 603593; -
DR MIM; 222700; -
DR InterPro; IPR002293; AA/rel_prmease1.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
KW Transport; Amino-acid transport; Transmembrane; Disease mutation.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 222 242 POTENTIAL.
FT TRANSMEM 259 279 POTENTIAL.
FT TRANSMEM 304 324 POTENTIAL.
FT TRANSMEM 383 403 POTENTIAL.
FT TRANSMEM 416 436 POTENTIAL.
FT TRANSMEM 441 461 POTENTIAL.
FT TRANSMEM 461 461 POTENTIAL.
FT CARBOHYD 54 54 G -> V (IN LPI).
FT VARIANT 54 54 L -> R (IN LPI).
FT VARIANT 334 334 /FTid=VAR_010261.
FT VARIANT 338 338 L -> R (IN LPI).
FT VARIANT 338 338 /FTid=VAR_010262.
FT VARIANT 386 386 G -> D (IN LPI).
FT VARIANT 386 386 /FTid=VAR_010999.
FT CONFLICT 91 91 S -> R (IN LPI).
FT CONFLICT 91 91 /FTid=VAR_011000.
FT CONFLICT 91 91 A -> V (IN REF. 4).
SQ SEQUENCE 511 AA; 55990 MW; A71D677B6B075894 CRC64;
Alignment Scores:
Pred. No.: 0.00109 Length: 511
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
Db: 1 Gaps: 0
US-09-667-170A-440 (1-2239) x YIAL_HUMAN (1-511)
QY 517 TATGCTGAATTCGGAACTATAAGAAATCTGGA 552
|||||
DB 90 TyrAlaGluLeuGlyThrThrIleLysSerGly 101
RESULT 4
RS21_AQUAE
ID RS21_AQUAE STANDARD; PRT; 67 AA.
AC O67028;
DT 30-MAY-2000 (Rel. 39, Created)

<p> RX MEDLINE=75082355; PubMed=1110754; RA Nakashima Y., Wiseman R.L., Konigsberg W., Marvin D.A.; RT "Primary structure and sidechain interactions of Pfl filamentous RL bacterial virus coat protein."; RN Nature 253:68-71(1975). [3] RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS). RX MEDLINE=94310463; PubMed=8036516; RA Liu D.J., Day L.A.; RT "Pfl virus structure: helical coat protein and DNA with paraxial RL phosphates."; RN Science 265:671-674(1994). [4] RP STRUCTURE BY NMR. RX MEDLINE=92022521; PubMed=1925543; RA Namudripad R., Stark W., Opella S.J., Makowski L.; RT "Membrane-mediated assembly of filamentous bacteriophage pfl coat RL protein."; RN Science 252:1305-1308(1991). CC -1- FUNCTION: COAT PROTEIN B IS THE MAJOR COAT PROTEIN OF THE VIRION CC -1- SUBCELLULAR LOCATION: Type I membrane protein. CC CC This SWISS-PROT entry is copyright. It is produced through a colla- CC boration between the Swiss Institute of Bioinformatics and the EMBL out- CC ribute by the European Bioinformatics Institute. There are no restriction CC use by non-profit institutions as long as its content is not modified CC and this statement is not removed. Usage by and for com- CC mercial entities requires a license agreement (See http://www.isb-sib.ch/) CC or send an email to license@isb-sib.ch. CC CC EMBL; X52107; CAA36331.1; - DR PIR; A04229; VCBPPF. DR PIR; S15143; S15143. DR PIR; S20695; S20699. DR PDB; 1IFM; 31-JUL-94. DR PDB; 1IFN; 31-JUL-94. DR PDB; 2IFM; 01-JAN-96. DR PDB; 2IFN; 01-JAN-96. DR PDB; 3IFM; 01-JAN-96. DR PDB; 4IFM; 01-JAN-96. DR PDB; 1PFI; 26-JAN-95. KW Coat protein; Signal; 3D-structure; Transmembrane. FT SIGNAL 1 36 FT CHAIN 37 82 COAT PROTEIN B. FT DOMAIN 37 55 EXTRACELLULAR. FT TRANSMEM 56 78 FT DOMAIN 79 82 INTRACELLULAR. FT CONFLICT 45 45 E -> Q (IN REF. 2). FT SEQUENCE 82 AA; 8377 MW; 25FD055DA4661504 CRC64; SQ </p>	<p> Alignment Scores: Pred. No.: 23-2 Length: 82 Score: 8.00 Matches: 8 Percent Similarity: 100.00% Conservative: 0 Best Local Similarity: 100.00% Mismatches: 0 Query Match: 1.09% Indels: 0 DB: 1 Gaps: 0 </p>
<p> US-09-667-170A-440 (1-2239) x COAB_BPPF1 (1-82) QY 1432 CTGGCAGTTCGTGGCGATTAT 1455 Dbb 69 LeuAlaValAlaGlyLeuIleTyr 76 RESULT 6 ID FAFR_ECOLI CD P23774; STANDARD; PRT: 260 AA. DT 01-NOV-1991 (Rel. 20, Created) DD 01-NOV-1995 (Rel. 32, Last sequence update) ED 01-NOV-1995 (Rel. 32, Last annotation update) GN 987P fimbrial operon positive regulatory protein fapR OS FAPR. </p>	

Hurwitz J.;
 Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 [3]
 SEQUENCE FROM N.A.
 Martindale D.W., Koop B.F.;
 Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: THE ELONGATION OF PRIMED DNA TEMPLATES BY DNA POLYMERASE
 DELTA AND EPSILON REQUIRES THE ACTION OF THE ACCESSORY PROTEINS
 PROLIFERATING CELL NUCLEAR ANTIGEN (PCNA) AND ACTIVATOR 1. THE
 40 kDa SUBUNIT BINDS ATP.
 -!- SUBUNIT: HETEROPENTAMER OF SUBUNITS OF 140/145, 40, 38, 37, AND
 36.5 kDa THAT FORMS A COMPLEX WITH PCNA IN THE PRESENCE OF ATP.
 -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 -!- DISEASE: Haploinsufficiency of RFC2 may be the cause of certain
 cardiovascular and musculo-skeletal abnormalities observed in
 Williams-Beuren syndrome (WBS), a rare developmental disorder. It
 is a contiguous gene deletion syndrome involving genes from
 chromosome band 7q11.23.
 -!- SIMILARITY: BELONGS TO THE ACTIVATOR 1 36 TO 40 kDa SUBUNITS
 FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; M87338; AAB09786.1; -.
 EMBL; AF045555; AAC04860.1; -.
 PIR; A42700; A42700.
 Genew; HGNC:9970; REC2.
 MIM; 600404; -.
 MIM; 194050; -.
 InterPro; IPR0033593; AAA_ATPase.
 InterPro; IPR003959; AAA_ATPase_centri.
 InterPro; IPR000862; RFCdomain.
 Pfam; PF00004; AAA; 1.
 SMART; SM00382; AAA; 1.
 DNA replication; ATP-binding; Nuclear protein;
 Williams-Beuren syndrome.
 NP_BIND 76 83
 FT CONFLICT 244 244 G -> L (IN REF. 1; AA SEQUENCE).
 SQ SEQUENCE 354 AA; 39157 MW; B50AC8EEF89F64A9 CRC64;

 Alignment Scores:
 Pred. No.: 18.5 Length: 354
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 1 Gaps: 0

 US-09-667-170A-440 (1-2239) x AC14_HUMAN (1-354)
 QY 1761 CTGCCCAAGGGGAGACACAAAATA 1784
 |||||||||||||||||||||||||
 Db 130 LeuProLysGlyArgHisLysIle 137

 RESULT 8
 AC14_CHICK
 ID AC14_CHICK STANDARD; PRT: 359 AA.
 AC P53033;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Activator 1 40 kDa subunit (Replication factor C 40 kDa subunit) (A1
 DE 40 kDa subunit) (RF-C 40 kDa subunit) (RFC40).
 GN RFC2.
 OS Gallus gallus (Chicken).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.
 ON NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Leghorn; TISSUE=Cochlea;
 RX MEDLINE=94347215; PubMed=8068208;
 RA Oberholtzer J., Cohen E.L., Davis J.G.;
 RT "Molecular cloning of a chick cochlea cDNA encoding a subunit of DNA
 RT replication factor C/activator 1";
 RL DNA Cell Biol. 13:857-863(1994).
 CC -!- FUNCTION: THE ELONGATION OF PRIMED DNA TEMPLATES BY DNA POLYMERASE
 CC DELTA AND EPSILON REQUIRES THE ACTION OF THE ACCESSORY PROTEINS
 CC PROLIFERATING CELL NUCLEAR ANTIGEN (PCNA) AND ACTIVATOR 1. THE
 CC 40 kDa SUBUNIT BINDS ATP (BY SIMILARITY).
 CC -!- SUBUNIT: HETEROPEPTAMER OF SUBUNITS OF 140/145, 40, 38, 37, AND
 CC 36.5 kDa THAT FORMS A COMPLEX WITH PCNA IN THE PRESENCE OF ATP (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: BELONGS TO THE ACTIVATOR 1 36 TO 40 kDa SUBUNITS
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U12438; AAA20552.1; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003595; AAA_ATPase_cent.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 KW DNA replication; ATP-binding; Nuclear protein.
 FT NP_BIND 81 88
 SQ SEQUENCE 359 AA; 39706 MW; A2FD38F964CD1F9 CRC64;
 Alignment Scores:
 Pred. No.: 18.4 Length: 359
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 1 Gaps: 0
 US-09-667-170A-440 (1-2239) x AC14_CHICK (1-359)
 QY 1761 CTGCCCCAAGGGAGACACACAAATA 1784
 Db 135 LeuProlGlyArgHisLysIle 142
 [1]
 RESULT 9
 GSC_DROME
 ID GSC_DROME STANDARD; PRT; 419 AA.
 AC P54366; Q9VPR9;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Homeobox protein goosecoid.
 GN GSC OR CG2851.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=96202483; PubMed=8625850;
 RA Cortely A., Stella M., Coffinier C., Kessler D., Mailhos C.,
 RA Dessain S., Desplan C.;

RT "A functional homologue of goosecoid in Drosophila.";
 RL Development 122:1641-1650(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96272167; PubMed=8670808;
 RA Hahn M., Jackie H.;
 RT "Drosophila goosecoid participates in neural development but not in
 RT body axis formation.";
 RL EMBO J. 15:3077-3084(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yeung M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassaman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 CC Science 287:2185-2195(2000).
 CC -!- FUNCTION: APPEARS TO REGULATE REGIONAL DEVELOPMENT OF SPECIFIC
 CC TISSUES. CAN RESCUE AXIS POLARITY IN UV-RADIATED XENOPUS
 CC EMBRYOS.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: IN EARLY EMBRYO DEVELOPMENT, EXPRESSION
 CC CONFINED TO TWO REGIONS: A HORSESHOE-LIKE PATTERN ACROSS THE
 CC DORSAL SIDE WHICH IS DESTINED TO FORM THE BRAIN HEMISPHERES AND A
 CC SECOND DOMAIN WHICH INVAGINATES INSIDE THE STOMODEUM AND WHICH, IS
 CC FATED TO FORM THE FOREGUT, RING GLAND AND STOMATOGASTRIC NERVOUS
 CC SYSTEM (SNS).
 CC -!- SIMILARITY: BELONGS TO THE PAIRED HOMEOBOX FAMILY.
 CC "BICOID" SUBFAMILY.
 CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-5 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

DR EMBL; X95420; CAA64699.1; -.
DR EMBL; U52968; AAB17948.1; -.
DR EMBL; AE003389; AAF51473.1; -.
DR HSSP; P06601; IFJL.
DR TRANSFAC; T04041; -.
DR Flybase; FBgn010323; Gsc.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Developmental protein; Nuclear protein; DNA-binding; Homeobox.
FT DOMAIN 104 107 POLY-ALA.
FT DOMAIN 164 169 POLY-SER.
FT DOMAIN 195 199 POLY-ALA.
FT DNA_BIND 286 345 HOMEBOX.
SQ SEQUENCE 419 AA; 44949 MW; 851A4C46AA861FB9 CRC64;

Alignment Scores:
Pred. No.: 18 Length: 419
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x GSC_DROME (1-419)
QY 1724 TCTCTTCGTFACAACTTCCAGT 1701
Db 370 SerSerSerGlyThrThrSerSer 377

RESULT 10
2A5E_HUMAN
ID 2A5E_HUMAN STANDARD; PRT; 467 AA.
AC Q16537;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit,
DE epsilon isoform (PP2A, B subunit, B' epsilon isoform) (PP2A, B
DE subunit, B56 epsilon isoform) (PP2A, B subunit, PR61 epsilon isoform)
DE (PP2A, B subunit, R5 epsilon isoform).
GN PP2R5E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 449-455.
RP TISSUE=Fetal retina;
RC MEDLINE=96276417; PubMed=8694763;
RA Zolnierowicz S., van Hoof C., Andjelkovic N., Cron P., Stevens I.,
RA Merlievede W., Goris J., Hemmings B.A.;
RT "The variable subunit associated with protein phosphatase 2A0 defines
RT a novel multimer family of regulatory subunits."
RL Biochem. J. 317:187-194(1996).
RN [2]
SEQUENCE FROM N.A.
RP TISSUE=Fetal brain;
RC MEDLINE=96355607; PubMed=8703017;
RX McCright B., Rivers A.M., Audlin S., Virshup D.M.;
RA "The B56 family of protein phosphatase 2A (PP2A) regulatory subunits
RT encodes differentiation-induced phosphoproteins that target PP2A to
RT both nucleus and cytoplasm."
RL J. Biol. Chem. 271:22081-22089(1996).
CC -1- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
CC COMPARTMENT.
CC -1- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
CC

```

CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS B (THE R2/B/PR55/B55, R3/B''/PR72/PR130/PR59 AND R5/B'/B56 FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS, AND CELL SIGNALING MOLECULES.

-1- SUBCELLULAR LOCATION: Cytoplasmic.

-1- PTM: PHOSPHORYLATED ON SERINE RESIDUES.

-1- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z69029; CAA93153.1; -
 EMBL; L76703; AAB69752.1; -
 Genew; HGNC:9313; PPP2R5E.
 MIM; 601647; -
 InterPro; IPR002554; B56.
 Pfam; PF01603; B56; 1.
 Phosphorylation; Multigene family.
 KW SEQUENCE 467 AA; 54699 MW; DD9CE11433F499CF CRC64;
 SQ SEQUENCE 467 AA; 54699 MW; DD9CE11433F499CF CRC64;

Alignment Scores:

Pred. No.:	17.7	Length:	467
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	1	Gaps:	0

US-09-667-170A-440 (1-2239) x 2A5E_HUMAN (1-467)

QY 79 AAGAAGAAAAAGACAGAGAGAA 102
 |||||
 Db 438 LysLysGluLysGluArgGluGlu 445

RESULT 11
 AMPA_CHLPN
 ID AMPA_CHLPN STANDARD; PRT; 499 AA.
 AC Q928F8; Q9JQJ2;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
 DE (LAP) (Leuyl aminopeptidase).
 GN PEPA OR CPN0385 OR CP0370.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CW1029;
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia

```

RT pneumoniae AR39. ";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=J138;
RA MEDLINE=20330349; PubMed=10871362;
RA Shira M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- FUNCTION: PRESUMABLY INVOLVED IN THE PROCESSING AND REGULAR
CC TURNOVER OF INTRACELLULAR PROTEINS. CATALYZES THE REMOVAL OF
CC UNSUBSTITUTED N-TERMINAL AMINO ACIDS FROM VARIOUS PEPTIDES (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-l-
CC xbb-, in which Xaa is preferably Leu, but may be other amino acids
CC including Pro although not Arg or Lys, and Xbb may be Pro.
CC -!- COFACTOR: MANGANESE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001623; AAD18529.1; -
DR EMBL; AE002199; AAF38219.1; -
DR EMBL; AP002546; BAA98593.1; -
DR HSSP; P00727; IBPN.
DR MEROPS; M17.UPW.
DR PHC1-2DPAGE; Q928F8; -
DR TIGR; CP0370; -
DR InterPro; IPR000819; Peptidase_M17.
DR Pfam; PF00883; Peptidase_M17; 1.
DR Pfam; PF02789; Peptidase_M17_N; 1.
DR PRINTS; PR00481; LAMNOPPTDASE.
DR PROSITE; PS00631; CYTOSOLAP; 1.
KW Hydrolase; Aminopeptidase; Manganese; Complete proteome.
FT METAL 263 263
FT METAL 268 268 MANGANESE 2 (BY SIMILARITY).
FT METAL 286 286 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 286 286 MANGANESE 2 (BY SIMILARITY).
FT METAL 345 345 MANGANESE 2 (BY SIMILARITY).
FT METAL 347 347 MANGANESE 1 (BY SIMILARITY).
FT ACT_SITE 275 275 MANGANESE 1 AND 2 (BY SIMILARITY).
FT ACT_SITE 349 349 POTENTIAL.
SQ SEQUENCE 499 AA; 54509 MW; B0F281B1DF4B9EC2 CRC64;

Alignment Scores:
Pred. No.: 17,5 Length: 499
Score: 8,00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 1 Gaps: 0

US-09-667-170a-440 (1-2239) x AMPA_CHLPN (1-499)
QY 380 GAGACTCCCTCAGTAACTGACT 357
Db 149 GluThrProLeuSerLysValThr 156

RESULT 12
LAT1_HUMAN
ID LAT1_HUMAN STANDARD; PR7; 507 AA.
AC Q01650; Q9UBN8; Q9UP15; Q9UQC0;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Large neutral amino acids transporter small subunit 1 (L-type amino
DE acid transporter 1) (4F2 light chain) (4F2LC) (4F2LC) (CD98 light
DE chain) (Integral membrane protein E16) (hLAT1).
GN SLC7A5 OR LAT1 OR MPE16 OR CD98LC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98421678; PubMed=9751058;
RA Mastroberardino L., Spindler B., Pfeiffer R., Skelly P.J., Loffing J.,
RA Shoemaker C.B., Verrey F.;
RT "Amino-acid transport by heterodimers of 4F2hc/CD98 and members of a
RT permease family."
RL Nature 395:288-291(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Ovary;
RA Yanagida O., Segawa H., Miyamoto K., Takeda E., Goya T., Endou H.,
RA Kanai Y.;
RT "Cloning and characterization of a human system L amino acid
RT transporter."
RN [3]
RP Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RX SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=99160855; PubMed=10049700;
RA Prasad P.D., Wang H., Huang W., Kekuda R., Rajan D.P., Leibach F.H.,
RA Ganapathy V.;
RT "Human LAT1, a subunit of system L amino acid transporter: molecular
RT cloning and transport function."
RL Biochem. Biophys. Res. Commun. 255:283-288(1999).
RN [4]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=99172172; PubMed=10072483;
RA Tsurudome M., Ito M., Takebayashi S., Okumura K., Nishio M.,
RA Kawano M., Kusagawa S., Komada H., Ito Y.;
RT "Primary structure of the light chain of fusion regulatory protein-
RT 1/CD98/4F2 predicts a protein with multiple transmembrane domains that
RT is almost identical to the amino acid transporter E16."
RL J. Immunol. 162:2462-2466(1999).
RN [5]
RP SEQUENCE FROM N.A.
RX Minato N., Iwai K., Takizawa C., Nakamura E.;
RT "Human 4F2 light chain: amino acid transporter."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 267-507 FROM N.A.
RX TISSUE=Periphereal blood lymphocytes;
RA Gaugitsch H.W., Prieschl E.E., Kalthoff F., Huber N.E.,
RA Baumrucker T.;
RT "A novel transiently expressed, integral membrane protein linked to
RT cell activation. Molecular cloning via the rapid degradation signal
RT AUUA."
RL J. Biol. Chem. 267:11267-11273(1992).
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE
CC NEUTRAL AMINO ACIDS INVOLVED IN CELLULAR AMINO ACID UPTAKE.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN ADULT LUNG AND LIVER,
CC AND IS ALSO EXPRESSED IN BRAIN, THYMUS, RETINA AND SOME OTHER
CC TISSUES.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

Thu Apr 17 09:48:33 2003

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

CC EMBL; AF077866; AAC61479.1; -
CC EMBL; AF018009; BAA84648.1; -
CC EMBL; AF104032; AAD20464.1; -
CC EMBL; AB018542; BAA33851.1; -
CC EMBL; AB017908; BAA75746.1; -
CC EMBL; AB0244; AAA35780.1; -
CC PIR; A42783; A42783; SLIC7A5.
CC Genew; HGNC:11063; SLIC7A5.
CC InterPro; IPR002293; AA/rel_primease1.
CC InterPro; IPR004760; L_AA_transport.
CC InterPro; IPR004841; Permease.
CC Pfam; PF00324; aa_permeases; 1.
CC TIGRFAMs; TIGR00911; 2A0308; 1.
CC TransPort; Amino-acid transport; Transmembrane.
CC TRANSMEM 50 70 POTENTIAL.
CC TRANSMEM 84 104 POTENTIAL.
CC TRANSMEM 120 140 POTENTIAL.
CC TRANSMEM 146 166 POTENTIAL.
CC TRANSMEM 170 190 POTENTIAL.
CC TRANSMEM 199 219 POTENTIAL.
CC TRANSMEM 243 263 POTENTIAL.
CC TRANSMEM 274 294 POTENTIAL.
CC TRANSMEM 319 339 POTENTIAL.
CC TRANSMEM 396 416 POTENTIAL.
CC TRANSMEM 431 451 POTENTIAL.
CC TRANSMEM 458 478 POTENTIAL.
CC CARBOHYD 49 49 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CONFLICT 15 15 A -> V (IN REF. 5).
CC CONFLICT 29 31 AKS -> SKR (IN REF. 5).
CC CONFLICT 35 35 S -> A (IN REF. 5).
CC CONFLICT 62 62 T -> A (IN REF. 5).
CC CONFLICT 88 88 V -> M (IN REF. 5).
CC CONFLICT 154 154 T -> A (IN REF. 5).
CC CONFLICT 230 230 N -> K (IN REF. 1).
CC SEQUENCE 507 AA; 55010 MW; 767F3C60B62C0F02 CRC64;

Alignment Scores:
Pred. No.: 17.5 Length: 507
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170a-440 (1-2239) x LAT1_HUMAN (1-507)

QY 517 TATGCTGAATGGGAACAATA 540
Db 103 TyrAlaGluLeuGlyThrThrille 110
|||||

RESULT 13
LAT1_MOUSE STANDARD; PRT; 512 AA.
AC Q92127; Q9JMI4; -
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Large neutral amino acids transporter small subunit 1 (L-type amino
DE acid transporter 1) (4F2 light chain) (4F2 LC) (4F2LC).
GN SLIC7A5 OR LAT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

STRAIN=BALB/C;
MEDLINE=99115648; PubMed=9915839;
RA Nakamura E., Sato M., Yang H., Miyagawa F., Harasaki M., Tomita K.,
RA Matsuoka S., Noma A., Iwai K., Minato N.;
RT "4F2 (CD98) heavy chain is associated covalently with an amino acid
RT transporter and controls intracellular trafficking and membrane
RT topology of 4F2 heterodimer.";
RL J. Biol. Chem. 274:3009-3016(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RA Kanai Y., Watanabe M., Endou H.;
RA "Localization of expression of system L neutral amino acid transporter
RA LAT1 in brain.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE
CC NEUTRAL AMINO ACIDS. INVOLVED IN CELLULAR AMINO ACID UPTAKE.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC TRANSPORT PROTEIN SLIC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AB017189; BAA75520.1; -
CC EMBL; AB023409; BAA90556.1; -
CC MGD; MGI:1298205; SLIC7a5.
CC InterPro; IPR002293; AA/rel_primease1.
CC InterPro; IPR004760; L_AA_transport.
CC InterPro; IPR004841; Permease.
CC Pfam; PF00324; aa_permeases; 1.
CC TIGRFAMs; TIGR00911; 2A0308; 1.
CC TransPort; Amino-acid transport; Transmembrane.
CC TRANSMEM 51 71 POTENTIAL.
CC TRANSMEM 85 105 POTENTIAL.
CC TRANSMEM 121 141 POTENTIAL.
CC TRANSMEM 147 167 POTENTIAL.
CC TRANSMEM 174 194 POTENTIAL.
CC TRANSMEM 200 220 POTENTIAL.
CC TRANSMEM 248 268 POTENTIAL.
CC TRANSMEM 279 299 POTENTIAL.
CC TRANSMEM 324 344 POTENTIAL.
CC TRANSMEM 398 418 POTENTIAL.
CC TRANSMEM 436 456 POTENTIAL.
CC TRANSMEM 463 483 POTENTIAL.
CC CARBOHYD 50 50 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 345 345 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CONFLICT 8 8 R -> M (IN REF. 2).
CC SEQUENCE 512 AA; 55902 MW; B3AB446E77374B0B CRC64;

Alignment Scores:
Pred. No.: 17.5 Length: 512
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170a-440 (1-2239) x LAT1_MOUSE (1-512)

QY 517 TATGCTGAATGGGAACAATA 540
Db 104 TyrAlaGluLeuGlyThrThrille 111
|||||

RESULT 14

```


LAT1_RAT
ID LAT1_RAT STANDARD; PRT; 512 AA.
AC O63016; O9QWL4;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Large neutral amino acids transporter small subunit 1 (L-type amino acid transporter 1) (4F2 light chain) (4F2 LC) (4F2LC) (Integral membrane protein E16) (TAL protein).
DE SLC7A5 OR MPE16 OR TAL.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98395066; PubMed=9726963;
RA Kanai Y., Segawa H., Miyamoto K., Uchino H., Takeda E., Endou H.;
RT "Expression cloning and characterization of a transporter for large neutral amino acids activated by the heavy chain of 4F2 antigen (CD98).";
RT J. Biol. Chem. 273:23629-23632(1998).
RN [2]
RP SEQUENCE OF 272-512 FROM N.A.
RC TISSUE=Hepatoma;
RX MEDLINE=95171385; PubMed=7532544;
RA Sang J., Lim Y.P., Panzica M., Finch P., Thompson N.L.;
RT "TAL, a highly conserved oncofetal complementary DNA from rat hepatoma, encodes an integral membrane protein associated with liver development, carcinogenesis, and cell activation.";
RT Cancer Res. 55:1152-1159(1995).
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE NEUTRAL AMINO ACIDS INVOLVED IN CELLULAR AMINO ACID UPTAKE.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: EXPRESSED HEPATOMA BUT NOT IN NORMAL LIVER.
CC ALSO EXPRESSED IN PLACENTA, TESTIS, BRAIN, OVARY, SPLEEN, MAMMARY GLAND, AND UTERUS.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT) SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AB015432; BAA33035.1; -
CC EMBL; U00995; AAA74411.1; -
CC InterPro; IPR002293; AA/rel_prmeasel.
CC InterPro; IPR004760; L_AA_transport.
CC Pfam; PF00324; aa_permeases; 1.
CC TIGRfams; TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 121 141 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 248 268 POTENTIAL.
FT TRANSMEM 279 299 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 398 418 POTENTIAL.
FT TRANSMEM 436 456 POTENTIAL.
FT TRANSMEM 463 483 POTENTIAL.
FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 512 AA; 55903 MW; B487CE0B58D73A02 CRC64;
Alignment Scores:
Pred. No.: 17.5 Length: 512
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0
US-09-667-170A-440 (1-2239) x LAT1_RAT (1-512)
QY 517 TATGCTGAATTGGACACACTATA 540
|||||
DB 104 TyrAlaGluLeuGlyThrThrIle 111
RESULT 15
LAT2_MOUSE
ID LAT2_MOUSE STANDARD; PRT; 531 AA.
AC Q9QXW9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Large neutral amino acids transporter small subunit 2 (L-type amino acid transporter 2).
DE SLC7A8 OR LAT2.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=20044753; PubMed=10574970;
RA Rossier G., Meier C., Bauch C., Summa V., Sordat B., Verrey F., Kuhn L.C.;
RT "LAT2, a new basolateral 4F2hc/CD98-associated amino acid transporter of kidney and intestine.";
RT J. Biol. Chem. 274:34948-34954(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20079165; PubMed=10610726;
RA Bassi M., Sperandio M.P., Incerti B., Bulfone A., Pepe A., Surace E.M., Gattuso C., de Grandi A., Buoinconti A., Riboni M., Manzoni M., Andria G., Ballabio A., Borsani G., Sebastio G.;
RT "SLC7A8, a gene mapping within the lysinuric protein intolerance critical region, encodes a new member of the glycoprotein-associated amino acid transporter family.";
RT Genomics 62:297-303(1999).
CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE NEUTRAL AMINO ACIDS. HAS HIGHER AFFINITY FOR L-PHENYLALANINE THAT LAYS A ROLE IN BASOLATERAL (RE)ABSORPTION OF NEUTRAL AMINO ACIDS.
CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: MAINLY EXPRESSED IN KIDNEY AND INTESTINE.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT) SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AF171668; AAF20380.1; -
CC EMBL; Y19022; CAB69072.1; -
CC MGI; I1355323; SLC7a8.
CC InterPro; IPR002293; AA/rel_prmeasel.
CC

DR InterPro; IPR004760; L_AA_transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMs; TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 39 59 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 230 250 POTENTIAL.
FT TRANSMEM 267 287 POTENTIAL.
FT TRANSMEM 309 329 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 387 407 POTENTIAL.
FT TRANSMEM 421 441 POTENTIAL.
FT TRANSMEM 446 466 POTENTIAL.
SQ SEQUENCE 531 AA: 57873 MW; AE9C3B42F3B24F8C CRC64;

Alignment Scores:
Pred. No.: 17.4 Length: 531
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x LAT2_MOUSE (1-531)

QY 1126 TCAATGTCAGTGGCAGTGACCTTT 1149
Db 298 SerAsnAlaValAlaValThrPhe 305
|||||

Search completed: April 16, 2003, 16:36:44
Job time : 33.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:27:06 ; Search time 96.5 Seconds
(without alignments)
9561.440 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 732

Sequence: 1 ggaggttgagtgagcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09667170/runat_04042003_091001_21199/app_query.fasta_1.2375
-DB=SPTREMBL_21 -QPWT=fastan -SUFFIX=olin2p.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170@cgn_1_132_@runat_04042003_091001_21199 -NCPD=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	ID	Description
1	446	60.9	494	4	Q9BYH2 homo sapien

ID	Q9BYH2	PRELIMINARY;	PRT;	494 AA.
AC	Q9BYH2;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Cystine/glutamate exchanger.			
GN	HXCT.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Kim J.Y., Chairoungdua A., Cha S.H., Segawa H., Matsuo H., Kim D.K.,			
RA	Endou H., Kanai Y.,			
RT	"Human cystine/glutamate exchanger: cDNA cloning and upregulation by			
RT	oxidative stress in glioma cells."			
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB040875; BAB40574.1; -			
DR	InterPro; IPR002293; AA/rel.prmeasel.			
DR	InterPro; IPR004760; L_AA_transp.			
DR	InterPro; IPR004841; Permease.			
DR	Pfam; PF00324; aa_permeases; 1.			

ALIGNMENTS

RESULT 1

ID	Q9BYH2	PRELIMINARY;	PRT;	494 AA.
AC	Q9BYH2;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Cystine/glutamate exchanger.			
GN	HXCT.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Kim J.Y., Chairoungdua A., Cha S.H., Segawa H., Matsuo H., Kim D.K.,			
RA	Endou H., Kanai Y.,			
RT	"Human cystine/glutamate exchanger: cDNA cloning and upregulation by			
RT	oxidative stress in glioma cells."			
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB040875; BAB40574.1; -			
DR	InterPro; IPR002293; AA/rel.prmeasel.			
DR	InterPro; IPR004760; L_AA_transp.			
DR	InterPro; IPR004841; Permease.			
DR	Pfam; PF00324; aa_permeases; 1.			

DR TIGRFAMS; TIGR00911; 2A0308; 1.
SQ SEQUENCE 494 AA; 54423 MW; BEC57A7EA4DDB0B1 CRC64;

Alignment Scores:

[illegible]

US-09-667-170A-440 (1-2239) X 09BYH2 (1-494)

QY	334	GAGAAAGTCAGCTGAAGAGGAAAGTCACCTACTTACTGAGGGGAGGTCTCCATTATCATCTGGC	393
DB	36	GLULysValGlnLeuLysArgLysValThrLeuLeuArgGlyValSerIleIleIleGly	55
QY	394	ACCATCATGGAGCAGGAATCTTCATCTCTCTAAGGCGGTGCTCCAGAACACGGGCAGC	453
DB	56	ThrIleIleGlyAlaGlyIlePheIleSerProLysGlyValLeuGlnAsnThrGlySer	75
QY	454	GTGGGCATGTCCTGACCATCTGGACGGGTGTGTGGGGTCCTGTCACATATTGGAGCGTTTG	513
DB	76	ValGlyMetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuPheGlyAlaLeu	95
QY	514	TCATTATCTCAATTGGGAACAACATAAAGAANAATCTGCAGCTCATACACATATATTTCG	573
DB	96	SerTyrAlaGluLeuGlyThrThrIleLysLysSerGlyGlyHisTyrThrTyrIleLeu	115
QY	574	GAAGTCCTTTGGTCCATTACCAGCTTTGTACAGAGTCTGGGTGGAACCTCCTCATATAATACGC	633
DB	116	GluValPheGlyProLeuProAlaPheValArgValTrpValGluLeuLeuIleIleArg	135
QY	634	CCTGCAGCTACTCCTGTGATATCCCTGGCATTTGGACGCTACATCTCGAACCATTTTTTT	693
DB	136	ProAlaAlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProPhePhe	155
QY	694	ATTCAATGTGAATCCCTGAACTTCGGAATCAAGCTCATACAGCTGTGGGCATAACGTGA	753
DB	156	IleGlnCysGluIleProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrVal	175
QY	754	GTGATGTCTCTAAATAGCATGAGTGTACGCTGGAGCGCCCGATCCAGATTTCTTAACC	813
DB	176	ValMetValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThr	195
QY	814	TTTTGCAAGCTCACAGCAATCTCGATAAATTATAGTCCCTGGAGTATGTCAGCTAAATAAA	873
DB	196	PheCysLysLeuThrAlaIleLeuIleIleIleValProGlyValMetGlnLeuIleLys	215
QY	874	GGTCAAAACGAGAACCTTTAAAGACGCGTTTTTCAGGAAGAGATTCAGATTTACCGGTTG	933
DB	216	GlyGlnThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerSerIleThrArgLeu	235
QY	934	CCACTGGCTTTTTTATTATGGAATGTATGATATGCTGGCTGGTTTTACCTCAACTTTGTT	993
DB	236	ProLeuAlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTrpPheTyrLeuAsnPheVal	255
QY	994	ACTGAAGAAGTAGAAAACCCCTGAAAAACCATTCGCCCTTGCAATATGTATATCCATGCCC	1053
DB	256	ThrGluGluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAla	275
QY	1054	ATTGTCCACCATTTGGCTATGTCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG	1113
DB	276	IleValThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrThrIleAsnAlaGlu	295
QY	1114	GAGTCGTGCTTTCAATGCAGTGGCAGTGACCTTTTCTGAGCGGCTACTCGGAAATTTTC	1173
DB	296	GluLeuLeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPhe	315
QY	1174	TCATTAGCAGTTCGGAFTTTGTTCCTGCTTTGGCTTCATGAACGGTGGTGTG	1233
DB	316	SerLeuAlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyVal	335

Qy	1234	TTTGCTCTCCAGGTTATCTATGTTGGCTCTCAGAGGGTCACCTTCAGAAATCCTC	129
Db	336	PheAlaValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIleLeu	355
Qy	1294	TCCATGATTTCATCGCGAAGACACCTCCTCTACCAAGCTGTTATTGTTTTCACCCCTTGG	1353
Db	356	SerMetIleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeu	375
Qy	1354	ACAAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTTGAATTCCTCAGTTTGGCC	1413
Db	376	ThrMetIleMetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPheLeuSerPheAla	395
Qy	1414	AGGTGGCTTTTATTGGCTGGCAGTGGCTGGCTGATTATCTTCGATACAAATGGCCA	1473
Db	396	ArgTrpLeuPheIleGlyLeuAlaValAlaGlyLeuIleTyrLeuArgTyrLysCysPro	415
Qy	1474	GATATGCATCGCTCTTTCAGGTGGCCACCTGTTTCATCCAGCTTGTGTTTTCCTTCACATGC	1533
Db	416	AspMetHisArgProPheLysValProLeuPheIleProAlaLeuPheSerPheThrCys	435
Qy	1534	CTCTTCATGGTGGCTTTCCTCTATTTCGGACCCCATTTAGTACAGGGAATTCGCTTCGTC	1593
Db	436	LeuPheMetValAlaLeuSerLeuTyrSerAspProPheSerThrGlyIleGlyPheVal	455
Qy	1594	ATCACTCTGACTGGAGTCCCTGGTATTATCTCTTTATTATATGGACAAGAAACCCAGG	1653
Db	456	IleThrLeuThrGlyValProAlaTyrTyrLeuPheIleIleTrpAspLysLysProArg	475
Qy	1654	TGGTTTAGAATAATGTCA	1671
Db	476	TrpPheArgIleMetSer	481

RESULT 2

Q9P1X0	PRELIMINARY;	PRT;	92 AA.
ID	Q9P1X0		
AC	Q9P1X0;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Cystine/glutamate transporter (Fragment).		
GN	XCT.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae;		
OX	NCBI_TaxID=9606;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RX	Medline=11213471;		
RA	Sato H., Tamba M., Kuriyama-Matsumura K., Okuno S.,		
RT	"Molecular cloning and expression of human xCT, the		
RT	amino acid transport system xc-";		
RL	Antioxid. Redox Signal. 2:1665-671(2000).		
DR	EMBL; AB042201; BAA94999.1; -.		
DR	InterPro; IPR002293; AA/re1_primeasel.		
FT	NON_TER	92	
SO	SEQUENCE	92 AA;	9676 MW; EA88DD0FE801B9CF CRC664

Alignment Scores:

Pred. No.:	5,73e-51	Length:	92
Score:	57.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	7.79%	Indels:	0
DB:	4	Gaps:	0

US-09-667-170A-440 (1-2239) x Q9P1X0 (1-92)

334 GAGAAGTGCAGCTCAAGCAGGAAAGTCACTTTACTAGGGAGCTCCATTATCATTTGGC 393
QY
36 GtuLysValGlnLeuLysArgLysValThrLeuLeuArgLysValSerIleIleIleGly 55
Db
394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGGCGTGTCCAGAACACGGGCAGC 453
QY

Db 56 ThrilleleGlyAlaGlyIlePheSerProLysGlyValLeuGlnAsnThrGlySer 75
 QY 454 GTGGCATGCTCTGACCATCTGACCGTGTGGGGCTCTGTCACATATT 504
 Db 76 ValGlyMetSerLeuThrIleTrpValCysGlyValLeuSerLeuPhe 92
 RESULT 3
 Q9QWS1
 ID Q9QWS1 PRELIMINARY; PRT; 510 AA.
 AC Q9QWS1
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Y+LAT1b.
 GN SLC7A7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=NIH/SWISS; TISSUE=HEART;
 RX MEDLINE=99094891; PubMed=9878049;
 RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;
 RT "Amino acid transport of y+L-type by heterodimers of 4F2hc/CD98 and
 RT members of the glycoprotein-associated amino acid transporter
 RT family.";
 RL EMBO J. 18:49-57(1999).
 DR EMBL; AJ130943; CAA10255.1;
 DR MGD; MGI:1337120; SLC7A7.
 DR InterPro: IPR002293; AA/rel_prmeasel.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 FT VARIANT 4 4
 FT SEQUENCE 510 AA; 55691 MW; 9F30FA7BAF176F6C CRC64;
 SQ SEQUENCE 510 AA; 55691 MW; 9F30FA7BAF176F6C CRC64;
 Alignment Scores:
 Pred. No.: 0.00362 Length: 510
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 11 Gaps: 0
 US-09-667-170A-440 (1-2239) x Q9QWS1 (1-510)
 QY 517 TATGCTGAATGGGAACAACATAAGAAATCTGGA 552
 Db 91 TyrAlaGluLeuGlyThrThrIleLysSerGly 102
 RESULT 4
 Q9Z1K8
 ID Q9Z1K8 PRELIMINARY; PRT; 510 AA.
 AC Q9Z1K8
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Y+LAT1a (Solute carrier family 7 (Cationic amino acid transporter, y+
 DE system), member 7).
 GN SLC7A7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=BALB/C; TISSUE=KIDNEY;
 RX MEDLINE=99094891; PubMed=9878049;
 RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;
 RT "Amino acid transport of y+L-type by heterodimers of 4F2hc/CD98 and
 RT members of the glycoprotein-associated amino acid transporter
 RT family.";
 RL EMBO J. 18:49-57(1999).
 Alignment Scores:
 Pred. No.: 0.00362 Length: 510
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 11 Gaps: 0
 US-09-667-170A-440 (1-2239) x Q9QWS1 (1-510)
 QY 517 TATGCTGAATGGGAACAACATAAGAAATCTGGA 552
 Db 91 TyrAlaGluLeuGlyThrThrIleLysSerGly 102
 RESULT 5
 Q9R0S5
 ID Q9R0S5 PRELIMINARY; PRT; 512 AA.
 AC Q9R0S5
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Y+L amino acid transporter 1.
 GN RY+LAT1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RC TISSUE=KIDNEY;
 RP SEQUENCE FROM N.A.
 RA Kanai Y., Fukasawa Y., Segawa H., Endou H.;
 RT "Characterization of a system y+L amino acid transporter.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB020520; BAA87325.1;
 DR InterPro: IPR002293; AA/rel_prmeasel.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 SQ SEQUENCE 512 AA; 55641 MW; 708905B2ACB0A824 CRC64;
 Alignment Scores:
 Pred. No.: 0.00362 Length: 512
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 11 Gaps: 0
 US-09-667-170A-440 (1-2239) x Q9R0S5 (1-512)
 QY 517 TATGCTGAATGGGAACAACATAAGAAATCTGGA 552
 Db 93 TyrAlaGluLeuGlyThrThrIleLysSerGly 104
 RESULT 6
 Q9QZ66
 ID Q9QZ66 PRELIMINARY; PRT; 512 AA.
 AC Q9QZ66
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)

[2]
 RN SEQUENCE FROM N.A.
 RP Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ012754; CAA10170.1;
 DR EMBL; BC014709; AAH14709.1;
 DR MGD; MGI:1337120; SLC7A7.
 DR InterPro: IPR002293; AA/rel_prmeasel.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 FT VARIANT 4 4
 FT SEQUENCE 510 AA; 55677 MW; 9F30FB1B88126F6C CRC64;
 SQ SEQUENCE 510 AA; 55677 MW; 9F30FB1B88126F6C CRC64;
 Alignment Scores:
 Pred. No.: 0.00362 Length: 510
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 11 Gaps: 0
 US-09-667-170A-440 (1-2239) x Q9Z1K8 (1-510)
 QY 517 TATGCTGAATGGGAACAACATAAGAAATCTGGA 552
 Db 91 TyrAlaGluLeuGlyThrThrIleLysSerGly 102
 RESULT 5
 Q9R0S5
 ID Q9R0S5 PRELIMINARY; PRT; 512 AA.
 AC Q9R0S5
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Y+L amino acid transporter 1.
 GN RY+LAT1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RC TISSUE=KIDNEY;
 RP SEQUENCE FROM N.A.
 RA Kanai Y., Fukasawa Y., Segawa H., Endou H.;
 RT "Characterization of a system y+L amino acid transporter.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB020520; BAA87325.1;
 DR InterPro: IPR002293; AA/rel_prmeasel.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 SQ SEQUENCE 512 AA; 55641 MW; 708905B2ACB0A824 CRC64;
 Alignment Scores:
 Pred. No.: 0.00362 Length: 512
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 11 Gaps: 0
 US-09-667-170A-440 (1-2239) x Q9R0S5 (1-512)
 QY 517 TATGCTGAATGGGAACAACATAAGAAATCTGGA 552
 Db 93 TyrAlaGluLeuGlyThrThrIleLysSerGly 104
 RESULT 6
 Q9QZ66
 ID Q9QZ66 PRELIMINARY; PRT; 512 AA.
 AC Q9QZ66
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)

DE Amino acid transporter y+LAT1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Ng A.M.L., Yao S.Y.M., Cheeseman C.I., Young J.D.;
 RT "cDNA encoding rat jejunal amino acid transporter y+LAT1.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF200684; AAF07216.1; -;
 DR InterPro: IPR002293; AA/rel_prmeasel.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 SQ SEQUENCE 512 AA; 55684 MW; 708905B2ACB2130B CRC64;

Alignment Scores:
 Pred. No.: 0.00362 Length: 512
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.64% Indels: 0
 DB: 11 Gaps: 0

US-09-667-170A-440 (1-2239) x Q90Z66 (1-512)

QY 517 TATGCTGAATTGGGAACACTATTAAGAAATCTGGA 552
 Db 93 TyrAlaGluLeuGlyThrIleLeuLysSergly 104
 |||||

RESULT 7

Q9NS82 ID Q9NS82 PRELIMINARY; PRT; 523 AA.
 AC Q9NS82;
 DT 01-OCT-2000 (TremBLrel. 15, Created)
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE Asc-type amino acid transporter 1 (ASCI protein).
 GN HASC-1 OR SLC7A10.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=20323545; PubMed=10863037;
 RA Nakauchi J., Matsuo H., Kim D.K., Goto A., Chairoungdua A., Cha S.H.,
 RA Inatomi J., Shiohara Y., Yamaguchi K., Saito I., Endou H., Kanai Y.;
 RT "Cloning and characterization of a human brain Na+-independent
 RT transporter for small neutral amino acids that transports D-serine
 RT with high affinity.";
 RT Neurosci. Lett. 287:231-235(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21400982; PubMed=11509015;
 RA Leclerc D., Wu Q., Ellis J.R., Goodyer P., Rozen R.;
 RT "Is the slc7a10 gene on chromosome 19 a candidate locus for
 RT cystinuria?"
 RT Mol. Genet. Metab. 73:333-339(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Bassi M.T.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Borsani G., Nunes V., Palacin M.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB037670; BAB03213.1; -;
 DR EMBL; AF340165; AAK93960.1; -;

DR EMBL; AF340155; AAK93960.1; JOINED.
 DR EMBL; AF340156; AAK93960.1; JOINED.
 DR EMBL; AF340157; AAK93960.1; JOINED.
 DR EMBL; AF340158; AAK93960.1; JOINED.
 DR EMBL; AF340159; AAK93960.1; JOINED.
 DR EMBL; AF340160; AAK93960.1; JOINED.
 DR EMBL; AF340161; AAK93960.1; JOINED.
 DR EMBL; AF340162; AAK93960.1; JOINED.
 DR EMBL; AF340163; AAK93960.1; JOINED.
 DR EMBL; AF340164; AAK93960.1; JOINED.
 DR EMBL; A277731; CAC81900.1; -;
 DR InterPro: IPR002293; AA/rel_prmeasel.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 SQ SEQUENCE 523 AA; 56797 MW; 24BA0B36521AC2D4 CRC64;

Alignment Scores:
 Pred. No.: 0.488 Length: 523
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.37% Indels: 0
 DB: 4 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9NS82 (1-523)

QY 409 GGAATCTTCATCTCTCTAAGGGCTGCTC 438
 Db 57 GlyIlePheIleSerProLysGlyValLeu 66
 |||||

RESULT 8

Q9JMH8 ID Q9JMH8 PRELIMINARY; PRT; 530 AA.
 AC Q9JMH8;
 DT 01-OCT-2000 (TremBLrel. 15, Created)
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE AscI (Asc-type amino acid transporter 1).
 GN SLC7A10 OR ASC-1.
 OS Mus musculus (Mouse), and
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090, 10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse;
 RX MEDLINE=20200467; PubMed=10734121;
 RA Fukasawa Y., Segawa H., Kim J.Y., Chairoungdua A., Kim D.K.,
 RA Matsuo H., Cha S.H., Endou H., Kanai Y.;
 RT "Identification and characterization of a Na+-independent neutral
 RT amino acid transporter that associates with the 4F2 heavy chain and
 RT exhibits substrate selectivity for small neutral d- and l- amino
 RT acids.";
 RT J. Biol. Chem. 275:9690-9698(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Rat; STRAIN=SPRAGUE DAWLEY; TISSUE=FOREBRAIN;
 RA Alberati-Giani D., Rew J.N.;
 RT "Cloning and characterization of a D-serine transporter from rat
 RT brain.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB026688; BAA93617.1; -;
 DR EMBL; AJ313387; CAC42885.1; -;
 DR MGI:1858261; SLC7A10.
 DR InterPro: IPR002293; AA/rel_prmeasel.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 SQ SEQUENCE 530 AA; 57548 MW; 0C5A80BF922DB54D CRC64;

Alignment Scores:
 Pred. No.: 0.486 Length: 530
 Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.37% Indels: 0
 DB: 11 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9JWH8 (1-530)

QY 409 GGAATCTTCATCTCTCTAGGGCGTCTC 438

Db 63 GlyllePheIleSerProLysglyValLeu 72

RESULT 9

Q9XMP1
 ID Q9XMP1 PRELIMINARY; PRT; 340 AA.
 AC Q9XMP1;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE NADH dehydrogenase (ubiquinone) chain 2 (EC 1.6.5.3).
 GN NDH-UI.
 OS Ceratitidis capitata (Mediterranean fruit fly).
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Tephritidae; Tephritidae; Ceratitidis.
 OC NCBI_TaxID=7213;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20225575; PubMed=10762421;
 RA Spanos L., Koutroumbas G., Kotsyfakis M., Louis C.;
 FT "the complete sequence of the mitochondrial genome of the Medfly,
 RT Ceratitidis capitata.";
 RL Insect Mol. Biol. 9:139-144(2000).
 CC -I- CATALYTIC ACTIVITY: NADH + UBIQUINONE - NAD(+) + UBIQUINOL.
 DR EMBL: AJ242872; CAB45088.1; ...
 DR InterPro: IPR001750; Oxidored.q1.
 DR Pfam: PF00361; oxidored.q1; 1.
 DR Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 KW SEQUENCE 340 AA; 39231 MW; F4C646E1CE5E3F4F CRC64;

Alignment Scores:

Pred. No.: 6.2 Length: 340

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.26% Indels: 0

DB: 8 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9XMP1 (1-340)

QY 2196 CTCCTCTTTCATATCATGATGAT 2170

Db 70 LeuLeuPheAlaIleIleMetMetTyr 78

RESULT 10

O57439
 ID O57439 PRELIMINARY; PRT; 507 AA.
 AC O57439;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE L amino acid transporter-1 LAT-1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95047611; PubMed=9829974;
 RA Torrents D., Estevez R., Pineda M., Fernandez E., Lloberas J.,
 RA Shi Y.B., Zorzano A., Palacin M.;
 RT "Identification and characterization of a membrane protein (y-L amino

RT acid transporter-1) that associates with 4F2hc to encode the amino
 RT acid transport activity y-L. A candidate gene for lysinuric protein
 RT intolerance.";
 RL J. Biol. Chem. 273:32437-32445(1998).
 DR EMBL: AF019906; AAB93541.1; ...
 DR InterPro: IPR002293; AA/rel_primeasel.
 DR InterPro: IPR004760; L_AA_transport.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 DR TIGRFAMs: TIGR00911; 2A0308; 1.
 DR TIGRFAMs: TIGR00911; 2A0308; 1.
 SQ SEQUENCE 507 AA; 55595 MW; EA0E9782B3313B82 CRC64;

Alignment Scores:

Pred. No.: 5.71 Length: 507

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.23% Indels: 0

DB: 13 Gaps: 0

US-09-667-170A-440 (1-2239) x O57439 (1-507)

QY 610 TGGTGGAAGCTCTCATATACGCCCT 636

Db 134 TrpValGluLeuLeuIleIleArgPro 142

RESULT 11

O13020
 ID O13020 PRELIMINARY; PRT; 507 AA.
 AC O13020;
 DT 01-JUL-1997 (TReMBLrel. 04, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Amino acid transporter chain (AmAt-L-1c) ASUR4.
 GN ASUR4.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97324256; PubMed=9178633;
 RA Spindler B., Mastroberardino L., Custer M., Verrey F.;
 RT "Characterization of early aldosterone-induced RNAs in A6 kidney
 RT epithelia.";
 RL Pflugers Arch. 434:323-331(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=98421678; PubMed=9751058;
 RA Mastroberardino L., Spindler B., Pfeiffer R., Skelly P.J., Loffing J.,
 RA Shoemaker C.B., Verrey F.;
 RT "Amino acid transport by heterodimers of 4F2hc/CD98 and members of a
 RT new permease family.";
 RL Nature 395:288-291(1998).
 DR EMBL: Y12716; CAA73254.1; ...
 DR InterPro: IPR002293; AA/rel_primeasel.
 DR InterPro: IPR004760; L_AA_transport.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 DR TIGRFAMs: TIGR00911; 2A0308; 1.
 DR CONFLICT 42 42 G -> A (IN REF. 0).
 FT CONFLICT 199 199 A -> R (IN REF. 0).
 FT CONFLICT 209 209 F -> L (IN REF. 0).
 FT CONFLICT 391 391 N -> D (IN REF. 0).
 SQ SEQUENCE 507 AA; 55529 MW; 2A6741CE89AB721D CRC64;

Alignment Scores:

Pred. No.: 6.2 Length: 340

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.26% Indels: 0

DB: 8 Gaps: 0

US-09-667-170A-440 (1-2239) x Q9XMP1 (1-340)

QY 2196 CTCCTCTTTCATATCATGATGAT 2170

Db 70 LeuLeuPheAlaIleIleMetMetTyr 78

RESULT 10

O57439
 ID O57439 PRELIMINARY; PRT; 507 AA.
 AC O57439;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE L amino acid transporter-1 LAT-1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95047611; PubMed=9829974;
 RA Torrents D., Estevez R., Pineda M., Fernandez E., Lloberas J.,
 RA Shi Y.B., Zorzano A., Palacin M.;
 RT "Identification and characterization of a membrane protein (y-L amino

Alignment Scores:

Pred. No.: 5.71 Length: 507

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

• [REDACTED]

100

RA Beja O., Koonin E.V., Aravind L., Taylor L.T., Seitz H., Stein J.L.,
 RA Bensen D.C., Feldman R.A., Swanson R.V., DeLong E.F.,
 RT "Comparative genomic analysis of coexisting archaeal genetic variants
 RT in an Antarctic marine microbial assemblage.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U40238; AAK66785.1; -.
 SQ SEQUENCE 89 AA; 9325 MW; B4E02F76DBE2F393 CRC64;

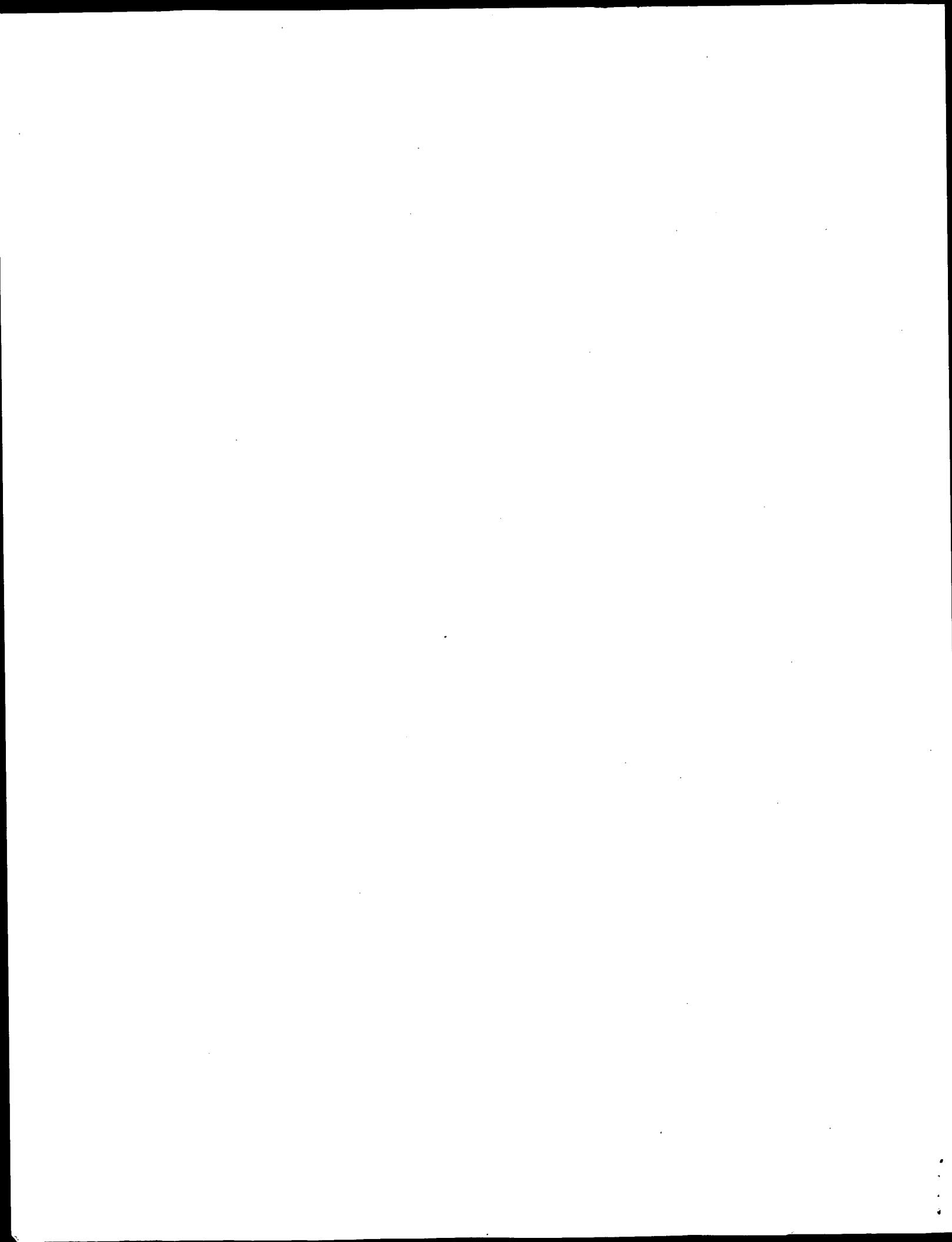
Alignment Scores:
 Pred. No.: 95.3 Length: 89
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.09% Indels: 0
 DB: 1 Gaps: 0

US-09-667-170A-440 (1-2239) x Q977T5 (1-89)

OY 385 ATCATGGCACCATCATGGAGCA 408

Db 19 ILeIlleGlyThrIleIleGlyAla 26

Search completed: April 16, 2003, 16:40:12
 Job time : 104.5 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:06:50 ; Search time 50 Seconds
(without alignments)
8609.799 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggsaggtgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q/Cgcn2_1/USPTO_spool/US09667170/runat_04042003_090918_20307/app_query.fasta_1.2375
-DB=PIR73 -OFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09667170 @CGN_1_1_57 @runat_04042003_090918_20307 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	1210.5	29.4	507	2 JG0165	LAT1 protein - hum
2	1058.5	25.7	493	2 T21445	hypothetical prote
3	975	23.7	563	2 T32479	hypothetical prote
4	910	22.1	464	2 T28818	hypothetical prote
5	715	17.4	566	2 T15226	hypothetical prote
6	672.5	16.3	662	2 T32821	hypothetical prote
7	665.5	16.2	541	2 T31554	hypothetical prote
8	622.5	15.1	562	2 T16854	hypothetical prote
9	564	13.7	440	2 T24837	hypothetical prote
10	521.5	12.7	438	2 B69855	amino acid permeas
11	479.5	11.6	440	2 E89921	hypothetical prote
12	463	11.2	455	2 A11995	hypothetical prote
13	404.5	9.8	424	2 B69172	cationic amino aci
14	368.5	8.9	574	2 S61943	methionine transpo

15	351.5	8.5	546	2	D87396	amino acid permeas
16	343.5	8.3	546	2	S48932	hypothetical prote
17	331	8.0	463	2	AE1155	amino acid transpo
18	331	8.0	583	2	T32266	hypothetical prote
19	330	8.0	503	2	T34694	probable cationic
20	328	8.0	463	2	AH1513	amino acid transpo
21	324	7.9	465	2	E69825	amino acid transpo
22	323	7.8	462	2	E65131	hypothetical 47.5
23	322	7.8	462	2	E91156	probable amino aci
24	322	7.8	462	2	B86002	probable amino aci
25	321.5	7.8	429	2	F71651	putrescine-ornithi
26	316.5	7.7	461	2	D69814	metabolite transpo
27	312.5	7.6	467	2	A97985	hypothetical prote
28	310.5	7.5	736	2	C69451	cationic amino aci
29	309.5	7.5	463	2	E95115	amino acid permeas
30	308.5	7.5	589	2	T32785	hypothetical prote
31	307.5	7.5	583	2	T48473	amino acid transpo
32	304	7.4	483	2	C82587	cationic amino aci
33	300.5	7.3	440	2	E70758	hypothetical prote
34	297	7.2	466	2	H97303	probable amino aci
35	295.5	7.2	427	2	F97790	putrescine-ornithi
36	295.5	7.2	441	2	C95307	probable transpo
37	295	7.2	776	2	A84178	cationic amino aci
38	294.5	7.1	463	2	AH1758	amino acid transpo
39	292.5	7.1	463	2	AE1383	amino acid transpo
40	292	7.1	465	2	C97751	cationic amino aci
41	289.5	7.0	447	2	G70030	amino acid permeas
42	289	7.0	438	2	F97234	ethanolamin permea
43	287.5	7.0	470	2	F71686	cationic amino aci
44	286	6.9	590	2	A86307	amino acid transpo
45	282.5	6.9	585	2	E88087	protein B0454.6 [1

ALIGNMENTS

RESULT 1

JG0165

LAT1 protein - human

C:Species: Homo sapiens (man)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000

C:Accession: JG0165; A42783

R:D.Prasad, P.; Wang, H.; Huang, W.; Kekuda, R.; P. Rajan, D.; H. Leibach, F.; Ganapath

Biochem. Biophys. Res. Commun. 255, 283-288, 1999

A:Title: Human LAT1, a subunit of system L amino acid transporter: Molecular cloning

A:Reference number: JG0165; MUID:99160855; PMID:10049700

A:Accession: JG0165

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-507 <DAP>

A:Cross-references: GB:AF104032; NID:94426639; PIDN:AAD20464.1; PID:94426640

R:Gaugitsch, H.W.; Prieschl, E.E.; Kaithoff, F.; Huber, N.E.; Baumruker, T.

J. Biol. Chem. 267, 11267-11273, 1992

A:Title: A novel transiently expressed, integral membrane protein linked to cell acti

A:Reference number: A42783; MUID:92283834; PMID:1597461

A:Accession: A42783

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 267-507 <GAU>

A:Cross-references: GB:M80244; NID:gi181907; PIDN:AAA35780.1; PID:gi181908

A:Experimental source: peripheral blood lymphocytes

A:Note: sequence extracted from NCBI backbone (NCBI:104749, NCBIP:104750)

C:Superfamily: arginine permease

C:Keywords: transmembrane protein

Alignment Scores:

Pred. No.:	6.65e-100	Length:	507
Score:	1210.50	Matches:	232
Percent Similarity:	66.19%	Conservative:	95
Best Local Similarity:	46.96%	Mismatches:	164
Query Match:	29.38%	Indels:	3
DB:	2	Gaps:	2

US-09-667-170A-440 (1-2239) x JG0165 (1-507)

QY 1321 CCTCTACACAGCTGTTATTTGTTTTCACCCCTTTGACAAATGATAATGCTCTCTCTGGAGAC 1380
Db 373 ProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 392
QY 1381 CTCGACAGCTCTTTGAATTTCTCAGTTTCCAGTTTGGCTGGCTGGAGTT 1440
Db 393 IlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIle 412
QY 1441 GCTGGCTGATTATCTTCGATACAAATGCCAGATATGTCATCGTCTCTTCAAGGTGCCA 1500
Db 413 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 432
QY 1501 CTTGTCATCCACAGCTTTGTTTCTTCATCCATCGCTCTTCATGGTGGCTTCCCTCTAT 1560
Db 433 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 452
QY 1561 TCGGACCCATTTAGTACAGGATTTGGCTTCGTCATCTGCTGAGTGGAGTCCCTGCGTAT 1620
Db 453 LysThrProValGluCysGlyIleGlyPheThrIleLeuSerGlyLeuProValTyr 472
QY 1621 TATCTCTTTTATTATATGGGACAAAGAACCCAGGTGGTTTAGAATAATCTCAGAGAAATA 1680
Db 473 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuGlnGlyIlePheSerThr 492
QY 1681 ACCAGAACATTAACAATAATCTGGAAGTTGTACCAGAGAA 1722
Db 493 ThrValLeuCysGlnLysLeuMetGlnValProGluGlu 506
RESULT 2
T21445
hypothetical protein F27C8.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T21445
R:LLoyd, C.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z19423
A:Accession: T21445
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-493 <WIL>
A:Cross-references: EMBL:Z68216; PIDN:CAA92459.1; GSPDB:GN00022; CESP:F27C8.1
C:Genetics:
A:Experimental source: clone F27C8
A:Gene: CESP:F27C8.1
A:Map position: 4
A:Introns: 60/1; 162/1; 204/1; 255/2; 385/3; 443/1; 474/1
C:Superfamily: arginine permease
Alignment Scores:
Pred. No.: 2,72e-86 Length: 493
Score: 1058.50 Matches: 202
Percent Similarity: 65.38% Conservative: 102
Best Local Similarity: 43.44% Mismatches: 156
Query Match: 25.69% Indels: 5
Gaps: 3
US-09-667-170A-440 (1-2239) x T21445 (1-493)

QY 346 CTGAAGAGAGAAAGTCACTTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCATTGGA 405
Db 28 LeuGluLysSerLeuThrLeuPheAsnGlyValSerMetIleValGlyCysIleIleGly 47
QY 406 GCAGGAATCTTCATCTCTCTCTTAAAGGGCTCTCCAGAACACAGCGGCGTGGCATGTCT 465
Db 48 SerGlyIlePheValSerProThrGlyValGlnGluGlnAlaGlySerValGlyLeuSer 67
QY 466 CTGACCATCTGGAGCGTGTGTGGGGCTCTCTCCTCATTTTGGAGCTTTGTCTTATGCTGAA 525
Db 68 LeuIleValTrpLeuIleSerGlyIlePheThrAlaIleGlyAlaTyrCysTyrAlaGlu 87
QY 526 TTGGGAACAACATATAAGAAATCTGGAGGTCATTCACATATATTTTGGAGTCTTTGGT 585
QY 247 CCTTCCTGGGCAACAGAGCCACCTGGCGAGAGCGCTTTTTCAGGAAGAGACGCCCTTT 306
Db 14 ProAlaAlaGluGluLysGluAlaArgGluLysMetLeuAlaAlaLysSerAlaAsp 33
QY 307 TCAGGAAGACAGCCCTTTTCAGGAAGAGAGAAAGTCCAGCTGAAGAGAAAGTCACTTTA 366
Db 34 GlySerAlaProAlaGlyGluGly--GluGlyValThrLeuGlnArgAsnIleThrLeu 52
QY 367 CTGAGGAGAGTCTCATTTATTCATTTGGCACCATTGGAGCAGGAATCTTCATCTCTCTCT 426
Db 53 LeuAsnGlyValAlaIleIleValGlyThrIleIleGlySerGlyIlePheValThrPro 72
QY 427 AAGGGCTGCTCCAGAACAGCGGCGTGGGATGCTCTCTGACCATCTGGAGCGGTGTGT 486
Db 73 ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValValTrpAlaAlaCys 92
QY 487 GGGCTCTGTCACATTTTGGAGCTTCTCTTATGCTCAATTTGGGAACAACATATAAGAAA 546
Db 93 GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrIleSerLys 112
QY 547 TCTGAGGTCAATTACACATATATTTTGGAGTCTTGGTCCATTACAGCTTTTGGTACGA 606
Db 113 SerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAlaPheLeuLys 132
QY 607 GTCGGGTGGAACTCCTCATATAATACGCCCTGACCTACTGCTGTGATATCCCTGGCATTT 666
Db 133 LeuTrpIleGluLeuIleAlaArgProSerSerGlnTyrIleValAlaLeuValPhe 152
QY 667 GGACGCTACATCTGGAACCACTTTTATTTCAATGTGAATCCCTGAACCTGGCATCAAG 726
Db 153 AlaThrTyrLeuLeuLysProLeuPheProThrCysProValProGluGluAlaAlaLys 172
QY 727 CTCATTACAGCTGGGCACTAATCTAGTGTCTCTTAATCATCATGATGCTGCTGCTGG 786
Db 173 LeuValAlaCysLeuCysValLeuLeuLeuThrAlaValAsnCysTyrSerValLysAla 192
QY 787 AGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTTCAGCAATTTCTGATAATATA 846
Db 193 AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuIleIle 212
QY 847 GTCCTGGAGTTATCGAGCTAATTAAGGTCAAGCGCAGAACTTTAAAGAGCCCTTTTCA 906
Db 213 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 232
QY 907 -----GGNAGAGATTCAGATTACGGGTGGCGCTGCTGCTGCTTTTATTTATGGAATGAT 960
Db 233 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 252
QY 961 GCATATGCTGCTGCTTTTACCTCAACTTTGTTACTGAAGAGTAGAAGAACCCCTGAAAAA 1020
Db 253 AlaTyrGlyTyrAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
QY 1021 ACATTCCTCTGCAATATGATATATCCATGCGCATTCACCATTTGGCTATGCTGAGCA 1080
Db 273 AsnLeuProLeuAlaIleIleIleSerLeuProIleValThrLeuValThrValLeuThr 292
QY 1081 AATGTGGCTACTTTACGACCACTTAAATGCTGAGAGCTGCTGCTTCAATTCAGTGGCA 1140
Db 293 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerGluAlaValAla 312
QY 1141 GTGACCTTTCTGAGCGGTACTGGGAATTTCTCATTTAGCATGTCCTGATCTTTGTTGCC 1200
Db 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 332
QY 1201 CTCTCTGCTTGGCTTCATGACGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPhePheVal 352
QY 1261 GCCTCTGAGAGGCTCACCTTCAGAAATCTCTCCATGATTTATGCTGCGGAGCACACT 1320
Db 353 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372

Db 88 LeuGlyThrLeuLeuLysSerGlyGlyAspTyrAlaTyrIleMetGluAlaPheGly 107
QY 586 CCATTACAGCTTTTGTACGAGCTGGTGGAACTCTCATATACGCCCTGAGCTACT 645
Db 108 ProPheValAlaPheIleArgLeuTrpIleGluAlaIleValValArgProCysThrVal 127
QY 646 GCTGTATATCCCTGGCAGCTTGGACGCTACATCTCGAACCACTTTTATTCAATGTGA 705
Db 128 ThrIleValAlaLeuThrPheAlaIleTyrGlyLeuArgProPheProAspCysAla 147
QY 706 ATCCCTGAATCGCATCAAGCTCATACAGCTGGGCGATCACTAGTGTGCTCTA 765
Db 148 ProProAspValAlaGluLeuLeuAlaIleLeuLeuIleValLeuMetThrAlaIle 167
QY 766 AATAGCATGATGTATAGTCCCTGGAGTATGCGCTAAATTAAGGTCAAAAGCAG 825
Db 168 AsnCysIleSerValArgLeuAlaThrIleValGlnAspThrIleAlaLysVal 187
QY 826 ACAGCAATCTGATATATAGTCCCTGGAGTATGCGCTAAATTAAGGTCAAAAGCAG 885
Db 188 ValAlaLeuCysLeuIleIleLeuThrGlyLeuGlyLeuLeuPhePheGlyGluSerGln 207
QY 886 -----AATTTAAAGACGCTTTTCAGGAAGAGATTCAGATTAATACGGGTGCCA 936
Db 208 TyrLysAspSerPheGluAsnIlePheGluAsnThrSerGlnAspPheThrLysValSer 227
QY 937 CTGCTTTTATATGGAATGATGCTATGCTGGTGTCTTACCTCAACTTTGTACT 996
Db 228 LeuAlaPheTyrSerGlyLeuPheAlaTyrSerGlyTyrPsnPheLeuAsnPheIleVal 247
QY 997 GAAGAAGTAGAAACCTCAAAACCACTCCCTGCAATATGATATATCATCCGCAAT 1056
Db 248 GluGluLeuGlnAsnProLysArgAsnLeuProLeuAlaIleAlaIleSerIleThrSer 267
QY 1057 GTCACCATGGCTATGCTGACAAATGTCGCTACTTTAGCAGCATTAATGCTGAGGAG 1116
Db 268 CysThrValIleTyrValLeuThrAsnValAlaLeuTyrThrAlaIleSerProAspGlu 287
QY 1117 CTGCTGCTTCAAAACGAGTGCAGTACCTTTTCTGAGCGGCTACTGGGAAATTTCTCA 1176
Db 288 MetLeuGluSerProAlaValAlaValLeuPheAlaAsnLysLeuTyrGlyLysPheAla 307
QY 1177 TTAGCAGTTCCGATCTTTGTGCTCTCTCCCTGCTTGGCTCCAGAACGCTGGTGT 1236
Db 308 PheIleMetProLeuCysValAlaCysSerThrIleGlySerAlaAsnGlyValIlePhe 327
QY 1237 GCTGCTCCAGGTATCTATGCTGCTCGAGAGGTCACCTTCAGAAATCCCTCC 1296
Db 328 ThrSerAlaArgLeuPheTyrSerGlyAlaArgGluGlyGlnMetProAlaValLeuThr 347
QY 1297 ATGATTCATGTCGCAAGCACACTCTCTACAGCTGTTATTTGTTGCAACCTTTGACA 1356
Db 348 MetIleAsnLysLysThrLysThrProIleProAlaValIleLeuThrGlyAlaLeuSer 367
QY 1357 ATG---ATAATGCTCTCTGAGAGACCTCGACAGCTTTTGAATTCCTCAGTTTGGC 1413
Db 368 IleAlaTyrLeuLeuAlaSerLysAspValTyrGlnLeuIleAsnTyrIleGlnIleSer 387
QY 1414 AGGTGGCTTTTATGTCGTCGAGTGTGCTGGCTGATTTATCTTCGATACAAATGCCA 1473
Db 388 TyrTrpLeuAlaIleGlyThrAlaIleAlaLeuPheTrpLeuArgThrMetPro 407
QY 1474 GATATGATCTCTTCAAGGTGCCACTGTTTCATCCAGCTGTGTTTCTTCACATGCC 1533
Db 408 AspAlaSerArgProIleLysValProLeuIleTrpProAlaIlePheLeuAlaGlyCys 427
QY 1534 CTCTTCATGCTGCTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1593
Db 428 IleAlaLeuValLeuIleProMetValAlaAlaProArgAspThrGlyIleGlyLeuLeu 447
QY 1594 ATCACTCTGACTGGAGTCCCTGGTATATCTCTTTATATATGAGCAAGAAACCCAGG 1653
Db 448 IleMetLeuSerAlaValProValTyrGlyIlePheIleGlyTrpLysAsnLysProLys 467

QY 1654 TGGTTTGAATAATGTCCAGAAAATAACCCAGAACATTACAAATAATACTGGAAGTTGTA 1713
Db 468 TrpPheAsnGluPheIleAspSerSerThrValPheIleGlnLysLeuPheMetValVal 487
QY 1714 CCAGAAGAAGATAAG 1728
Db 488 ---AspGluAspLys 491
RESULT 3
T32479
hypothetical protein F52H2.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Jun-2000
C:Accession: T32479
R:Chisoe, S.; Hawkins, J.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid F52H2.
A:Reference number: Z21175
A:Accession: T32479
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-563 <CHI>
A:Cross-references: EMBL:AF026214; PIDN:AAB71312.1; GSPDB:GN000028; CESP:F52H2.2
A:Experimental source: strain Bristol N2; clone F52H2
C:Genetics:
A:Gene: CESP:F52H2.2
A:Map position: X
A:Introns: 20/3; 60/1; 123/2; 162/1; 204/1; 255/2; 296/3; 329/1; 401/2; 433/3; 544/1
C:Superfamily: arginine permease
Alignment Scores:
Pred. No.: 8.36e-79 Length: 563
Score: 975.00 Matches: 190
Percent Similarity: 55.85% Conservative: 106
Best Local Similarity: 35.85% Mismatches: 160
Query Match: 23.67% Indels: 74
DB: 2 Gaps: 3
US-09-667-170A-440 (1-2239) x T32479 (1-563)

QY 346 CTGAAGAGAAAGTACACTTTACTGAGGGAGTCTCATTATCATTTGGCACCACATCTGGA 405
Db 28 LeuGluLysThrMetThrLeuPheAsnGlyValSerIleIleValGlyCysIleIleGly 47
QY 406 GCAGGAATCTTCATCTCTCTTAAGGGCGTCTCCAGAACACGGGCGAGCGTGGCATGTCT 465
Db 48 SerGlyIlePheIleSerProThrGlyIleGlnAlaGlnAlaGlySerValGlyLeuSer 67
QY 466 CTGACCATCTCGACGGTGTGGGGTCTCTGCTACTATTTGGAGCTTTGTCTTATGCTGAA 525
Db 68 LeuIleValTrpValLeuSerGlyLeuPheAlaGlyIleGlyAlaPheCysTyrAlaGlu 87
QY 526 TTGGACACACTATAAAGAAATCTGGAGGTCAATTACACATATATTTTGGAAAGTCTTTGT 585
Db 88 LeuGlyThrLeuIleArgLysSerGlyGlyAspTyrIleMetGluAlaPheGly 107
QY 586 CCATTACAGCTTTTGTACGAGTCTGGTGGAACTCTCATATATACGCCCTGAGCTACT 645
Db 108 ProPheLeuAlaPheLeuArgLeuTrpIleGluSerIleValValArgProCysThrAla 127
QY 646 GCTGTATATCCCTGGCATTTGGACGCTACATCTCGAACCACTTTTATTCAATGTGAA 705
Db 128 ThrIleValAlaLeuThrPheAlaIleTyrMetLeuLysProPheThrProAspCysAsp 147
QY 706 ATCCCTGAATCGCATCAAGCTCATACAGCTGGGCGATCACTAGTGTGCTGCTCA 765
Db 148 SerProLeuSerThrGluLeuIleAlaAlaLeuLeuLeuValLeuThrAlaVal 167
QY 766 AATAGCATGATGTACGCTGAGCGGCCCGGATCCAGATTTTCTTAACCTTTTGGCAAGCTC 825
Db 168 AsnCysIleSerValLysTrpAlaSerLysValGlnAspPhePheValThrLysThr 187

QY 826 ACAGCAATCTGATATTTAGTCCCTGGAGTATGACAGTAAATTAAGGTCAACGACG 885
Db 188 AlaAlaLeuValLeuIlePheThrGlyLeuTrpAsnMetValSerGlyLysProGlu 207
QY 886 -----AATTTAAAGACGCCCTTTTCAGGACAGAGATTCAGGTATTACGCGGTGGCCA 936
Db 208 AlaPheAspSerPheGluAsnIlePheGluAsnThrAlaLysAspLeuGluThrAlaSer 227
QY 937 CTGCTTTTATTTATGATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
Db 228 LeuAlaPheThrSerGlyLeuPheAlaTrpGlnGlyTrpAsnTyLeuAsnPheIleVal 247
QY 997 GAAGAAGTGAAGAACCTGAAAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
Db 248 GluGluLeuGlnAsnProLysArgAsnLeuProLeuSerIleAlaIleSerCysSerLeu 267
QY 1057 GTCACCATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
Db 268 CysThrIleIleThrLeuThrAsnValAlaLeuThrSerIleThrProAspGlu 287
QY 1117 CTGCTGCTTCAATGAGTGGCAGTGCACCTTTCTGAGCGGCTACTGGGAAATTTCTCA 1176
Db 288 MetLeuAlaSerProAlaValAlaValLeuPheAlaGluLysAsnTrpGlyTrpPheAla 307
QY 1177 TTACGAGTTCGGATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1233
Db 308 PheCysMetProIlePheValAlaCysSerThrIleGlySerAlaAsnGlyValIleLeu 327
QY 1233 ----- 1233
Db 328 ThrArgLeuCysGlySerArgAspValArgGlnAlaAlaIleAlaGluValAlaGluPro 347
QY 1233 ----- 1233
Db 348 GlnAsnProLysIlePheGlyAsnArgGlnLeuProAsnLeuThrSerGlnLysIlePhe 367
QY 1233 ----- 1233
Db 368 GlyAsnArgGlnLeuProPhePheGluAsnPhePheValGluIleIlePheLeuSerAla 387
QY 1234 -----TTTCCTGCTCCAGGTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1266
Db 388 LeuGluLysSerMetValPhePheThrSerGlnThrLeuPheThrCysGlyAla 407
QY 1267 CGAGAGGTTCATCCAGAAATCTCTCCATGATTCGCTGCTGCTGCTGCTGCTGCTGCT 1326
Db 408 ArgGluGlnMetProAsnValLeuThrMetValAsnLysThrThrLysThrProIle 427
QY 1327 CCAGCTGTTATGTTTGCACCTTTGACAAATGATA---ATGCTCTTCTGCTGAGACCTC 1383
Db 428 ProAlaValIleLeuThrGlyLeuLeuSerLeuLeuThrLeuLeuSerAsnAsnIle 447
QY 1384 GACAGTCTTTGAATTTCTCAGTTTCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1443
Db 448 TyrSerLeuIleAsnTrpIleGlnValSerThrTrpIleAlaIleGlyGlyAlaIleLeu 467
QY 1444 GGCTGCTTATTCGATACAAATGCCAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1503
Db 468 AlaLeuPheThrPheArgLysThrMetProAspAlaProArgAlaValLysAlaProIle 487
QY 1504 TCCATCCAGCTTTGTTTCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1563
Db 488 ValPheProIlePhePheIleGlyCysValLeuLeuValLeuValProValLeuGly 507
QY 1564 GACCATTTAGTACAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1623
Db 508 AsnProLysAspThrAlaIleGlyIleLeuIleMetLeuSerGlyValProValTrpLeu 527
QY 1624 CTCTTTATATATGAGACAAAGAACCCAGTGGTTAGTATATGCTGAGAAATATACC 1683
Db 528 IlePheIleAlaTrpLysGlyLysProLysCysIleAspSerLeuThrAspSerValThr 547
QY 1684 AGACATTTACAAATATATCTGGAAGTGTGA 1713

Db 548 IlePheThrGlnLysLeuPheMetValVal 557
RESULT 4
T28818
hypothetical protein F07C3.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
C:Accession: T28818
R:Favella, A.; Gattlung, S.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of C. elegans cosmid F07C3.
A:Reference number: Z20528
A:Accession: T28818
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-464 <FAV>
A:Cross-references: EMBL:U50308; PIDN:AAC48006.1; GSPDB:GN00023; CESP:F07C3.7
A:Experimental source: strain Bristol N2; clone F07C3
C:Genetics:
A:Gene: CESP:F07C3.7
A:Map position: 5
A:Introns: 91/3; 240/2; 272/3; 299/2
C:Superfamily: arginine permease
Alignment Scores: 5.37e-73 Length: 464
Pred. No.: 910.00 Matches: 185
Score: 60.81% Conservative: 99
Percent Similarity: 39.61% Mismatches: 159
Best Local Similarity: 22.09% Indels: 24
Query Match: 2 Gaps: 5
DB:
US-09-667-170A-440 (1-2239) x T28818 (1-464)
QY 337 AAGTCGAGCTGAAGAGAAATGACCTTTACTGAGGGAGTCTCCATTATATATGTCAC 396
Db 12 GlnIleLysLeuLysProArgIleSerLeuPheAsnGlyCysThrIleIleGlyVal 31
QY 397 ATCATTTGAGGAGAAATCTTCATCTCTCTTAAGGGCTGCTCCAGAACACGCGGACGCTG 456
Db 32 IleIleGlySerGlyIlePheValSerProLysGlyValLeuLeuGluAlaGlySerAla 51
QY 457 GGCATGCTCTGCACCATCTGGAGCTGTGTGGGTCTCTGTCATATTTGGAGCTTTGCT 516
Db 52 GlyMetSerLeuLeuIleTrpLeuLeuSerGlyValPheAlaMetIleGlyAlaValCys 71
QY 517 TATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGCTCATTTACACATATATTTGGAA 576
Db 72 TyrSerGluLeuGlyThrLeuIleProLysSerGlyAspTrpAlaTrpIleThrGlu 91
QY 577 GTCCTTTGCTGCTTACACAGCTTTGTAGAGTCTGGGTGGAACCTCTCATATATACGCT 636
Db 92 AlaPheGlyProLeuProSerPheLeuPheLeuTrpValAlaLeuValIleAsnPro 111
QY 637 CGAGCTACTGCTGATATCCCTGGCATTTGGAGCTGATCATCTGGAACCATTTTATT 696
Db 112 ThrSerLeuAlaIleAlaIleThrCysAlaThrTrpAlaLeuGlnProPheTrp 130
QY 697 CAATGTGAATCCCTGAACCTCCGATCAGCTCATTACAGCTGTGGGCAATCTAGT 756
Db 131 SerCysProValProAspValValValAsnLeuPheAlaGlyCysIleIleAlaValLeu 150
QY 757 ATGGCTTAATAGCATGATGCTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTT 816
Db 151 ThrPheIleAsnCysTrpAspValArgMetAlaThrArgThrAsnAspPheThrIle 170
QY 817 TGCAGCTCAGCAATCTGATATTTATAGTCCCTGAGTATGACAGTAAATTAAGGT 876
Db 171 ThrLysLeuIleAlaLeuThrLeuIleThrCysGlyGlyTrpLeuSerLeuGly 190
QY 877 CAACCCCAACAC-----TTTAAAGACGCTTTTTCAGGAGAGATTTCAAGTATTACCGG 930

Thu Apr 17 09:48:26 2003

```

305 PheSerArgGluThrLeuGlyAspPheAlaAsnValProPheLeuIleGlyLeu 324
1207 TGCTTTGGTCCATGAACGGTGTGTTGTCTCTCCAGTTATTCTATGTTGCGTCT 1266
325 LeuIleGlySerLeuAsnSerAsnLeuPheSerGlySerArgTyrMetTyrAlaAla 344
1267 CGAGAGGTCACCTCCAGAAATCCCTCCATGATTCGCGCAACACACCTCTCTA 1326
345 ArgGlnGlyHisLeuProAlaCysPheSerCysValAsnThrGluThrGluSerProArg 364
1327 CCAGCTGTTATTGTTTGCACCTTTGACAAATGATATGCTCTCTCTCGAGACCTCGAC 1386
365 ValAlaValLeuAlaGlnSerValLeuAlaValLeuSerTyrIleGlyAspLeuAsp 384
1387 AGTCTTTTGAATTCCTCCAGTTTGGCAGGTGCTTTTATTGGGCTGCGCAGTTGCTGG 1446
385 ThrLeuIleThrTyrValMetPheGlyPheTrpAlaGlnArgIlePheSerLeuValAla 404
1447 CTGATTATTCGATACAAATGCCAGATATGATGAT---CGCTCTTCAAGTGCCACATG 1503
405 LeuLeuIleAlaArgHisAsnHisIleProValHisProAspAlaValArgValProLeu 424
1504 TTCATCCAGCTTCTTCTTCTCCATGCTCTCTCATGCTTCCCTCTCTCTCTCTCTATCG 1563
425 PheCysIleTyrLeuPheLeuAlaIleThrValAlaLeuValIleIleProIlePheTyr 444
1564 GACCATTTAGTACAGGATGTTGGTTCGTCATCATCTCTGACTGAGTCCCTGCGTATTAT 1623
445 GluPheGlnSerThrAlaLeuAlaIleAlaIleCysLeuPheGlyPheValLeuTyrTyr 464
1624 CTCCTTTATATATGACCAAGAACCCAGGTGTTAGATAATGTCAGAGAAATAAAC 1683
465 ValPheIleHisLysAlaIlePheProArgTrpLeuValAlaLeuAsnLysLysValThr 484
1684 AGAACATTACAATAATCTGGAAGTCTGACCAAGAAAGATAAGTTATGAATAATGGA 1743
485 LeuTrpCysIleLeuPheAspCysLeuProAspVal-----LysGlyGly 500
1744 CTTGAGATCTTGCCA 1758
501 ValGlnLeuLeuAla 505

```

RESULT 6

```

T32821
hypothetical protein F54D12.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32821
R:Maggi, L.; Goela, D.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid F54D12.
A:Reference number: Z21228
A:Accession: T32821
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-662 <MAG>
A:Cross-references: EMBL:AF040647; PTDN:AA94996.1; GSPDB:GN00020; CESP:F54D12.3
A:Experimental source: strain Bristol N2; clone F54D12
C:Genetics:
A:Gene: CESP:F54D12.3
A:Map position: 2
A:Introns: 31/1; 88/2; 191/1; 235/1; 304/1; 396/2; 443/3; 510/3; 584/3

Alignment Scores:
Pred. No.: 1,06e-51 Length: 662
Score: 672.50 Matches: 178
Percent Similarity: 43.01% Conservative: 96
Best Local Similarity: 27.94% Mismatches: 186
Query Match: 16.32% Indels: 177
Gaps: 2
DB: 13

```

US-09-667-170A-440 (1-2239) x T32821 (1-662)

```

337 AAAGTCCAGCTGAAGAGAAAGTCACTTTACTGAGGGGAGTCTCCATTATCATTTGGCACC 396
7 GluIleProGlnLysHisThrIleGlyLeuIleThrAlaIleSerTyrThrValGlyAsp 26
397 ATCATTTGGAGCAGGAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 456
27 IleValGlySerGlyIlePheIleSerProThrSerIleLeuAsnHisAlaGlySerVal 46
457 GGCATCTCTCTGACCATCTGGACGCTGTGTGGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 516
47 GlyLeuSerLeuCysLeuTrpAlaLeuCysAlaCysIleSerLeuPheGlyAlaLeuSer 66
517 TATCTGCAATTTGGGAACAATACTATAAGAAATCTGGAGGTCTTACACATATATT----- 570
67 TyrValGlnLeuGlyThrSerIleArgLysSerGlyCysAspPheAlaTyrLeuSerHis 86
570 ----- 570
87 PheGlySerThrGlnThrLeuAsnThrLeuArgGlnAsnAlaGlnIleLeuValIleCys 106
570 ----- 570
107 GluLysLysCysValGluLeuLysPheThrLeuAspPheThrGlnAsnSerGluLysAsn 126
571 -----TTGGAA 576
127 PheIleIlePhePheArgLysPheMetSerPheArgLeuLysLysLeuArgPheLeuLys 146
577 GTCCTTGGTCCATTA----- 591
147 SerPheThrProIlePheLeuThrPheIleThrGlnAsnValThrLysIleAlaIleIle 166
591 ----- 591
167 SerGlnPhePheHisThrLysAsnAlaThrLysSerHisArgThrHisPheSerValIle 186
592 -----CCAGCTTTTGTACGA----- 606
187 ThrLysAsnAlaLysAsnProAspPheProLysAsnIlePheArgArgProLeuAlaSer 206
607 -----GTCCTGGTGAACCTCTCATATAAGCCCTCAGCTACTGCTGTGATATCCCTG 660
207 SerPheMetTrpValSerThrCysLeuSerTyrProAlaValLeuAlaIleGlnAlaIle 226
661 GCATTTGGAGCTACATCTCTGGAACA----- 687
227 SerPheGlyGluTyrIleValThrGluLysThrLysIleLeuLysLysLysLysPhe 246
688 -----TTTTTTTCAATGTGAATCCCTGAACTGCGATCGATCAAGCTC 729
247 PheSerLysPheLeuAsnPheLeuPheLysTyrThrValSerTyrPheIleLeuLysIle 266
730 -----ATTACA----- 735
267 SerPhePheLysHisLeuIleIleLysPhePheGlnGlyLeuAspSerTrpIleThrIle 286
736 -----GCTGTGGGCATAAATGTA-----GTG 756
287 AspGluAsnTrpArgPheMetThrTyrArgLeuValGlyPheSerMetLeuTrpProLeu 306
757 ATGGTCTCTAAATAGCATGAGTGTACAGCTGAGCGCC---CGGATCCAGATTTTCTTAACC 813
307 MetLeuLeuAsnPhePheSerLeuLysValAlaGlyAlaPheGlnIleValAlaThr 326
814 TTTTGGCAAGCTCACAGCAATCTCTGATAATTATAGTCCCTGGAGTTATGACGCTAATT--- 870
327 AlaIleLysLeuIleValAlaAlaSerIleIleIleIleIleIleIleIleIleIlePhe 346
871 AAAGGTCAAACGCAAGACTTTTAAAGCGCTTTTTCAGAGAGAGATTCAAGATTACGGGG 930
347 LysGlnGlnThrGlnAsnPheLysAsnSerPheThrGlySerAspTrpAsnProGlyAsn 366

```


QY 931 TTGCCACTGGCTTTTATTATGGAATGATGATGATGATGCTGCTGCTTTTACTCAACTTT 990
 Db 367 LeuValLeuGlyValTyrSerGlyLeuPheAlaTyrAsnGlyTyrPaspValLeuAsnPhe 386
 QY 991 GTTACTGAAGAAGTAGAAGACCTGAAAGAACCATTCCTCCCTGCAATATGATATCCCATG 1050
 Db 387 GlyAlaGluGluGluAsnProArgArgThrLeuProIleAlaAlaIleSerGlyIle 406
 QY 1051 GCCATTGTGACCATTTGGCTGATGCTGACAAATGTGGCTACTTACGACCATTAATGCT 1110
 Db 407 AlaIleSerAlaThrValPheIleLeuMetAsnValSerTyrPheSerValLeuSerVal 426
 QY 1111 GAGGAGCTGCTGCTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152
 Db 427 GluAspPheLysAsnSerProAlaValAlaValAlaValArgThrPhePheIleThrPheAla 446
 QY 1153 GAGCGCTACTGGGAAATTTCTCAATAGCAGTTCGATTTTGTGGCTCTCTGCTGCTT 1212
 Db 447 GluArgThrLeuGlyAspPheHisTyrAlaIleProPheLeuIleSerLeuLeuLeu 466
 QY 1213 GCTCCATGAACGGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
 Db 467 GlySerMetAsnThrThrIlePheAlaCysSerArgTyrMetTyrSerGlyAlaGlnGln 486
 QY 1273 GGTCACTTCAGAAATCCTCTCCATGATTCATGCTCGCAAGCACACTCTCTACCAAGCT 1332
 Db 487 SerValMetProThrProLeuArgGlyIleHisArgThrArgSerProArgLeuAla 506
 QY 1333 GTTATTGTTTGCACCTTTGCAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
 Db 507 ValPheAlaGluIleLeuIleAlaIleCysLeuSerPheIleGlyAsnLeuAspGlnLeu 526
 QY 1393 TTGAATTTCTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1452
 Db 527 IleSerTyrMetSerPheAlaLeuTyrSerGlnArgThrCysThrGlnValGlyPheIle 546
 QY 1453 TATCTTCA-----TACAAAGCCAGATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Db 547 TyrPheLysLeuArgGlyThrLeuLysThrGlnAsp-----SerPheGlnValPro 563
 QY 1501 CTGTATCTCCAGCTTTGTTTCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 Db 564 IlePheValProValValPheLeuGlyIleCysIleAlaLeuValIleProIleThr 583
 QY 1561 TCGACCACTTAGTACAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
 Db 584 GlnAsnTyrHisValAlaIleTyrGlyValSerMetThrIleGlyAlaIleIleTyr 603
 QY 1621 TATCTCTTTATATATGGCAAGAACCCAGGTGTTAGATAATATGTCAGAGAAATA 1680
 Db 604 LeuIlePheIlePheProAsnThrLeuProIlePheLeuHisLysIleAsnSerIle 623
 QY 1681 ACCAGAACATACAAATATCTGGAAGTTGTA-----CCAGAAAGAT 1725
 Db 624 ValIysPheAlaGlnIleIlePheAsnCysValIleGluProTyrGluAsp 640

RESULT 7

T31554
 hypothetical protein Y53H1C.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T31554
 R:White, S.
 submitted to the EMBL Data Library, October 1999
 A:Reference number: Z21044
 A:Accession: T31554
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-541 <WIL>
 A:Cross-references: EMBL:AL117201; PIDN:CAB55063.1; CESP:Y53H1C.1
 A:Experimental source: clone Y53H1C
 C:Genetics:
 A:Gene: CESP:Y53H1C.1

A: Introns: 31/1; 90/3; 127/1; 154/1; 198/1; 249/2; 325/2; 371/2; 433/3; 467/1; 510/1
 Alignment Scores:
 Pred. No.: 4, 35e-51 Length: 541
 Score: 665.50 Matches: 156
 Percent Similarity: 55.93% Conservative: 108
 Best Local Similarity: 33.05% Mismatches: 169
 Query Match: 16.15% Indels: 39
 DB: 2 Gaps: 9

US-09-667-170A-440 (1-2239) x T31554 (1-541)

QY 376 GTCTCCATTATCATGTCACCATCATTTGGAGCAGGAATCTTCATCTCTCCTAAGGCGTG 435
 Db 26 MetSerTyrValIleAlaAsnIleIleGlyAlaGlyIlePheIleThrProGlyThrVal 45
 QY 436 CTCAGAACACGGCGCTGGCATGCTCTCAGCATCTGACCATCTGGAGCGTGTGGGCTGCTG 495
 Db 46 LeuLysLeuAlaLeuThrAsnGlyMetAlaLeuValTrpLeuGlyCysGlyLeuIle 65
 QY 496 TCACTATTTGGAGCTTTGCTTTATGCTGAATTTGGGAACAACATAAAGAAATCTGAGGT 555
 Db 66 SerLeuIleGlyIleCysTyrIleGluLeuGlyThrSerIleArgaspProGlyCys 85
 QY 556 CATTACACATATATTTTGGAGCTTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTG 615
 Db 86 AspPheAlaTyrAsnValTyrValGlyTrpGluGlyIleAlaPheSerPheMetTrpVal 105
 QY 616 GAACCTCTCATATACGCTCGCAGCTACTGCTGATATCCTCGCATTTTGGACGCTAC 675
 Db 106 GlyValIleMetSerPheProAlaSerAlaAlaValGlnAlaGlnThrPheGlyGlnTyr 125
 QY 676 ATCTGTG-----GAACCATTTTATTTCAATGTGA 705
 Db 126 IleValAlaGlyLeuSerProIleTrpHisLeuGlySerProTyrAspValIleLeuGlu 145
 QY 706 ATCCCTGAACCTTCCGATCAGCTCATTACAGCTGGGCGCATACGTAGTGTGCTTA 765
 Db 146 -----ArgGlyLeuGlyPheAlaLeuIleIleLeu 156
 QY 766 AATAGCATGAGTGT-----AGCTGGAGCGCCGGATCCAGATTTTCTTAACC 813
 Db 157 ThrValLeuAsnLeuTyrAlaIleAspLysTyrAlaSerLysPheGlnIlePheValThr 176
 QY 814 TTTTCAAGCTCACAGCAATCTGTATATATAGTCCCTGGAGTATGACGATAAT--- 870
 Db 177 IleAlaLysLeuSerLeuAlaIleIleIleValThrGlyPheTrpTyrLeuIleVal 196
 QY 871 AAAGTCAAAACGACAGAACTTTAAAGACGCTTTTCA-----GGAAGAGATTCAAGT 921
 Db 197 LysGlyGluThrGluHisPheLysAspAlaPheThrProLeuProAsnGluLysTyrAsp 216
 QY 922 ATTACGCGTTCCTGCTGCTGCTTTTATGGAATGTATGATCATGCTGCTGCTGCTTTTAC 981
 Db 217 IleGlyGlnIleSerLeuAlaPheTyrGlyAlaLeuTyrSerPheAlaGlyTrpAspIle 236
 QY 982 CTCAACTTTGTTACTGAAGAACCTGAGAACCTGAGAACCTGAGAACCTGAGAACCTGAG 1041
 Db 237 LeuAsnTyrGlyThrProGluIleLysAsnProArgThrMetProIleAlaLeuLeu 256
 QY 1042 ATATCCATGGCCATTTGCTCACCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
 Db 257 GlyGlyIleSerAlaValThrLeuValTyrMetAlaMetAsnValSerTyrMetThrVal 276
 QY 1102 ATTATGCTGAGGAGCTGCTGCTTTCAATGCAAGTGGAGCGTACCTTTTCTGACGGCTA 1161
 Db 277 LeuAspThrGluThrLeuLysAsnSerSerAlaValAlaAlaAspPheAlaArgIleThr 296
 QY 1162 CTGGAAATTTCTCATTAGCAGTTCCGATCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1221
 Db 297 LeuGlyAspPheSerTyrAlaIleProPheMetIleLeuLeuIleGlyThrLeu 316
 QY 1222 AACGGTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1281

```

Db 317 AsnSerAsnIlePheCysGlySerArgPheThrHisAlaAlaArgGluGlyHisLeu 336
Qy 1282 CCAGAAATCTCTCCATGATTCATGTCGCGAAGCAGCAGCTCTCTACAGCTGTTATGTT 1341
Db 337 ProThrPheLeuSerCysIleAsnAlaGluSerAsnSerProArgAlaAlaLeuLeuPhe 356
Qy 1342 TTGCACCTTTGACAATGATGCTCTCTCGGAGACCTCGACAGCTCTTTGAATTC 1401
Db 357 GlnLeuIleCysThrIleAlaValThrPheVal---AspThrGluSerLeuIleThrTyr 375
Qy 1402 CTCAGTTTGGCAGGTGCTTTTATTGGCTGGCAGCTCTGGCTGATTTATCTTCGA 1461
Db 376 ValThrPheValMetPheGlyGlnArgValPheThrMetAlaAlaLeuLeuTyrPheArg 395
Qy 1462 TACAAATGCCAGATATGCAT---CGTCCTTCAAGTGGCAGCTGTCATCCAGCTTTG 1518
Db 396 TyrArgAsnIleProValHisProAspAlaIleArgValProLeuIlePheSerIleLeu 415
Qy 1519 TTTTCTTCACATGCTCTCTCATGTTGCTGCTTTCCTCTATTTCGGACCCATTT 1572
Db 416 Phe-----PheMetIleThrIleAlaLeuValThrProPheIleGlu 430
Qy 1573 -----AGTACAGGATTTGGCTTCGTCATCATCTCTGAGTGGAGTCCCTGCTATAT 1623
Db 431 AspPheThrThrIleValGlyValGlyLeuValLeuMetGlyPheLeuLeuTyrMet 450
Qy 1624 CTCCTTTATATATGGGACAAACCCAGGTGTTAGATAATGTCAGAGAAATACCC 1683
Db 451 IlePheMetLysProLysGlnLeuProGlyPheLeuTyrArgPheAsnAspGlyIleThr 470
Qy 1684 AGAATATTACAATAATACTGGAAGTTGTCACAGAA 1719
Db 471 ArgValThrCysLysIleLeuPheThrThrProAsp 482

RESULT 8
T16854
hypoetical protein T13A10.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16854
R:Nelson, J.
submitted to The EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid T13A10.
A:Reference number: Z18590
A:Accession: T16854
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-562 <NEM>
A:Cross-references: EMBL:U56963; NID:g1293813; PID:g1293821; PIDN:AA38125.1; GSPDB:GN00
A:Experimental source: strain Bristol N2; clone T13A10
C:Genetics:
A:Gene: CESP-T13A10.10
A:Map position: 4
A:Introns: 23/2; 107/3; 140/1; 167/1; 211/1; 259/2; 335/2; 381/2; 500/1; 540/1

Alignment Scores:
Pred. No.: 3,11e-47 Length: 562
Score: 622.50 Matches: 154
Percent Similarity: 54.03% Conservative: 101
Best Local Similarity: 32.63% Mismatches: 180
Query Match: 15.11% Indels: 37
DB: 2 Gaps: 9

US-09-667-170A-440 (1-2239) x T16854 (1-562)
Qy 397 ATCATTGGAGCAGGATCTTCATCTCTCTAGGCGCTCTCCAGCAGCAGCGGCGTG 456
Db 50 IleIleGlyAlaGlyIlePheIleThrProGlyProIleLeuGlnIleThrPheSerAsn 69
Qy 457 GGATGCTCTGACCATCTGAGCGGTGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
Db 70 GlyLeuAlaLeuLeuValThrIleGlyCysGlyLeuIleSerLeuIleGlyIleCys 89

```

```

Qy 517 TATGCTGAATTTGGGAACAATATAAGAAATCTGGAGTCTATTACATATATATTTGAA 576
Db 90 TyrIleGluLeuGlyThrSerIleHisAspProGlyCysAspPheAlaTyrThrValTyr 109
Qy 577 GTCCTTTGGTCCATTAACAGCTCTTTTACAGTCTGGTGGGAGTCTCTCATATAGCCCT 636
Db 110 ValGlyTyrGluGlyIleAlaPheSerPheMetIleThrValGlyValIleMetSerPhePro 129
Qy 637 GCAGCTACTGCTGCTGATATCCCTGCTGCTTT-----GGAGCTATCATCTCTGGAACATTT 690
Db 130 AlaSerAlaAlaValGlnAlaLeuThrPheValAlaGlyMetAlaProIleThrProLeu 149
Qy 691 TTTTATTCAATGTGAATCCCTGAACCTTGGATTCGATCAAGCTCATTTACAGTGTGGGCTACT 750
Db 150 -----GluHisProThrAspGlyIle-----IleGluLysGlyLeuGlyPheAla 164
Qy 751 GTAGTATGCTGCTTAATAAGCATGAGTGTCT-----AGCTGGAGCCCGGATC 798
Db 165 LeuIleValLeuThrIleLeuAsnLeuTyrAlaIleAspLysIleAlaSerLysPhe 184
Qy 799 CAGATTTTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATATATATATATATATAT 858
Db 185 GlnIleValThrIleAlaLysMetLeuSerLeuAlaIleIleIleValThrGlyPhe 204
Qy 859 ATGACGCTAATT---AAAGTCAAAGCAGCAACTTTAAAGACGCTTTTTCAGGAAGAT 915
Db 205 TyrTyrLeuIlePheLysGlyGlnThrGluTyrLeuGluHisProPheGluGlySerAsn 224
Qy 916 TCAAGTATTACGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 975
Db 225 ThrAsnProGlyGlnIleSerLeuAlaPheTyrGlyAlaLeuThrPheAlaGlyTyr 244
Qy 976 TTTTACCTCAACTTTGTTTACTGAAGAAGTAGAAGAACCTGAAACCAATCCCTTCCCTTGA 1035
Db 245 AsnIleLeuAsnPheGlyThrProGluIleArgAsnProArgThrMetProIleAla 264
Qy 1036 ATATGATATATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1095
Db 265 LeuLeuGlyGlyValSerIleValThrAlaValTyrLeuAlaMetAsnIleSerTyrMet 284
Qy 1096 ACACCATTAATGCTGAGGAGTGTCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGAG 1155
Db 285 ThrValLeuThrProGlyGlnIleMetAsnSerThrAlaValAlaAlaAlaAspPheAlaGln 304
Qy 1156 CGCTACTGGGAAATTTCTCATTTAGCAGTTCGATCTTGTGCTGCTGCTGCTGCTGCTGCTG 1215
Db 305 IleThrLeuGlyGlyPheSerTyrAlaIleProPheMetIleAlaLeuLeuLeuGly 324
Qy 1216 TCATGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1275
Db 325 ThrLeuAsnSerAsnIlePheCysGlySerArgPheThrHisAlaAlaAlaArgGluGly 344
Qy 1276 CACCTCCGAAATCTCTCCATGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1335
Db 345 HisLeuProThrPheLeuSerCysIleAsnGluLysSerAsnSerProArgAlaAlaLeu 364
Qy 1336 ATTGTTTGGACCTTTGCAATGATATGCTCTCTCTGAGACCTGACAGCTCTTTTG 1395
Db 365 LeuPheGlnLeuValCysThrValValThrPheIle---AspThrAsnSerLeuIle 383
Qy 1396 AATTTCTCTGAGTTTGGCAGTGTGCTTTTATTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1455
Db 384 AsnTyrValSerPheValMetPheGlyGlnArgValPheThrMetThrAlaLeuMetTyr 403
Qy 1456 CTTGATACAAATGCCAGATATGAT---CGTCTTCTCAAGTGGCAGTGTTCATCCCCA 1512
Db 404 IleArgTyrArgAsnIleProValHisProAspAlaIleArgValProLeuIlePheSer 423
Qy 1513 GCTTTGTTTCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1548
Db 424 PheLeuPheThrProLeuIleThrIleAlaLeuValValProPheIleGluGluPheThr 443

```

QY 1549 -----CTTCTCCCTCTATTTCGGAC 1566
 Db 444 ValGlyValThrAspLeuProAsnLeuGluLysPheGlnAsnValGlnLeuLeuSerGlu 463
 QY 1567 CCATTT---AGPACAGGATTGGCTTCGTCTCATCTGAGTCCCTGGCTATTAT 1623
 Db 464 LysPheGlnThrIleValGlyValGlyLeuValLeuMetGlyValPheLeuTyrlle 483
 QY 1624 CTTCTATTATATGGCAAGAACCCAGGTGGTTTGAATAATGTCCAGAGAAATAACC 1683
 Db 484 IlePheMetLysProMetLysLeuProGluPheLeuIleArgPheAsnAspSerMetThr 503
 QY 1684 AGAACAATTACAATAATACTGCAAGTTGTACCCAGAA 1719
 Db 504 ArgIleThrCysLysIleLeuPheThrProAsp 515

RESULT 9

T24837

hypothetical protein T11F9.4 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T24837

R:Lennard, N.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z19941

A:Accession: T24837

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-440 <WTL>

A:Cross-references: EMBL:Z74042; PIDN:CAA98529.1; GSPDB:GN00023; CESP:T11F9.4

C:Genetics:

A:Gene: CESP:T11F9.4

A:Map position: 5

A:Introns: 41/3; 237/2; 269/3; 386/1

Alignment Scores:

Pred. No.:	5.18e-42	Length:	440
Score:	564.00	Matches:	136
Percent Similarity:	48.60%	Conservative:	90
Best Local Similarity:	29.25%	Mismatches:	159
Query Match:	13.69%	Indels:	80
DB:	2	Gaps:	10

US-09-667-170A-440 (1-2239) x T24837 (1-440)

QY 352 AGGAAGTCACCTTACTAGGGGAGTCTCCATTATGTCACCATTCATTGGAGCAGGA 411
 Db 8 GlnLysMetGlyLeuLeuGlyAlaIleSerTyrlleValGlyAsnIleValGlySerGly 27
 QY 412 ATCTTCATCTCTCTAAGGCGTGTCCAGACACGGGAGCGGTGGGCATCTCTGACC 471
 Db 28 IlePheIleThrProThrSerIleIleGluAsnValAsnSerIle----- 42
 QY 472 ATCTGGACGGTGTGTGGGGTCTGTGCATPATTGAGGCTTTCTCTATGCTGAATGGGA 531
 Db 43 -----GlnIleGlnGlyAla-----Gly 48
 QY 532 ACAACTATAAGAAATCTCGAGGTCAATACACATATATTTTGGAGTCTTTGGTCCATTA 591
 Db 49 ValGluLeuAspGluThrSer----- 55
 QY 592 CCAGCTTTTGTACGAGTCTGGGTGGGAACTCTCATAATACGCCCTGCGAGTACTGCTGTG 651
 Db 56 -----GluPheIlePheLysLysLeuLeu-----GlyPheSerLeuIle 68
 QY 652 ATATCCCTGGCATTTGGAGCTACATCTCTGGAACCATTTTATTCAATGTGAATCCCT 711
 Db 69 IleLeuLeuMetPheMetAsnPhePheSerLeuLysThrPheValGln----- 84
 QY 712 GAACCTGGCATCAAGCTCATACGCTGTGGGCATAACTGTGATGTGCTTAATAGC 771
 Db 84 ----- 84

QY 772 ATGAGTGTGAGTGGAGCGCCGGATCCAGATTTCCTTAACCTTTTCCCAAGCTCACAGCA 831
 Db 85 -----ArgPheSerIleLeuAlaSerLeuAlaLysIleAlaAa 97
 QY 832 ATTCGTATAATATAGTCCCTCGAGATTATGACGCTAATT---AAAGTCAAACGCAGAAC 888
 Db 98 ThrLeuLeuIleIleIleThrGlyPheTyrlleLeuIlePheLysHisThrLysGlnAsn 117
 QY 889 TTAAAGACGCTTTTTCAGGAGAGATTCAAGTATTACGCGGTGGCCACCTGGCTTTTAT 948
 Db 118 LeuGluGluProPheLysGlySerAsnTrpAsnProGlyProPheValAsnAlaLeuPhe 137
 QY 949 TATGGAATGTATGATGCTGGCTGCTTTTACCTCAACTTTGTTACTGAGGAAGTAGAA 1008
 Db 138 AlaGlyLeuPheSerTyrlleAspGlyTrpAspIleLeuAsnPheGlyAlaGluIleGlu 157
 QY 1009 AACCTGGAACCAATTCCTTGGCAATATGATATATCCATGCCATTGTACCATTTGGC 1068
 Db 158 AsnProLysArgThrMetProLeuSerIleIleGlyMetThrCysIleGlyValIle 177
 QY 1069 TATGCTGTGACAAATGGCTTACTTTACGACCATTAATGCTGAGGAGTGTGCTTTCA 1128
 Db 178 TyrValAlaValAsnValAlaTyrlleSerIleValLeuSerProThrGluMetIleAlaSer 197
 QY 1129 AATGAGTGGCAGTACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCG 1188
 Db 198 AsnAlaValAlaIleAspPheAlaAsnLysThrLeuGlyAlaAlaPheValValPro 217
 QY 1189 ATCTTTCTGCTCCCTCTCTCTGCTGCTCCATGCAAGCGGTGGTGTCTGCTCTCCAGG 1248
 Db 218 ValMetValAlaIleLeuLeuIleGlySerLeuAsnSerThrMetPheSerAlaSerArg 237
 QY 1249 TTATCTATGCTGCTGCTGAGAGGCTACCTTCCAGAAATCCTCTCATGATGCTCATGTC 1308
 Db 238 TyrLeuGlnAlaValSerArgGlnGlyHisIleProSerAlaIleSerGlyIleAlaPro 257
 QY 1309 CGCAAGCACACTCTCTACGAGTGTATTTGTTGACCCCTTTGACAAATGATAATGCTC 1368
 Db 258 AsnCysAspSerProArgValAlaLeuLeuValHisIleLeuIleAlaIleAlaValSer 277
 QY 1369 TTCTCTGGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTCCAGGTGGCTTTTATT 1428
 Db 278 PheLeuGlyAspProAspLysLeuIleAsnTyrlleValAlaPheAlaGlnTrpSerGlnArg 297
 QY 1429 GGCTGGCAGTGTGCTGGCTGATTTATCTCGATACAAATGCCAGATATGATCGTCT 1488
 Db 298 AlaPheThrMetSerAlaLeuLeuTyrlleArgIleArgGlyArgProArgHisProAsp 317
 QY 1489 ---TTCAAGTGGCAGTGTCTCCAGCTTTGTTTCTTCATGCTTCCATGCTCTTCATGTT 1545
 Db 318 ArgIleGlnLeuProIleIleMetProIleLeuPhePheLeuValCysThrSerMetVal 337
 QY 1546 GCCTTTCCCTCTATTTCGACCCCATTTAGTACAGGATTTGGCTTCGTCACTCACTGACT 1605
 Db 338 ValIleSerIleIleAspAspPheLysSerSerAlaValGlyLeuGlyIleLeuLeuGly 357
 QY 1606 GGAGTCCCTCGCTATTATCTCTTTTATTATATGGGACAAAG-----AAACCCAGGTGG 1656
 Db 358 GlyLeuIleIlePheIleIlePhe---ValTrpAspArgAlaLeuProSerSerHisThr 376
 QY 1657 TTAGA-----ATAATGTCAGAAATAAACCAAGCAATACAAATAACTACTG 1704
 Db 377 PheArgAsnAlaThrHisValIleAsnGluSerThrLysPheMetGlnIleIlePhe 396
 QY 1705 GAAGTTGTACCAAGAA 1719
 Db 397 AsnValValProGlu 401

RESULT 10

B69855

amino acid permease homolog ykba - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C::Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C::Accession: B69855

C::Runst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
R.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A::Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Ho
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshihida, K
A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A::Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A::Reference number: A69580; MUID:98044033; PMID:9384377

A::Accession: B69855

A::Status: preliminary; nucleic acid sequence not shown; translation not shown

A::Molecule type: DNA

A::Residues: 1-438 <KUN>

A::Cross-references: GB::Z99110; GB::AL009126; NID::Z62633472; PIDN::CABI3143.1; PID::Z6263640

A::Experimental source: strain 168

C::Genetics:

A::Gene: ykba

C::Superfamily: arginine permease

```

Alignment Scores:
Pred. No.:      3,310-38      Length:      438
Score:          521.50        Matches:     136
Percent Similarity: 50.46%    Conservative: 83
Best Local Similarity: 31.34% Mismatches:    196
Query Match:    12.66%       Indels:      19
DB:             2            Gaps:        8

US-09-667-170A-440 (1-2239) x B69855 (1-438)

```

Alignment Scores:

Pred. No.: 1.92e-34 Length: 440
 Score: 479.50 Matches: 125
 Percent Similarity: 49.33% Conservative: 96
 Best Local Similarity: 27.90% Mismatches: 194
 Query Match: 11.64% Indels: 33
 DB: 2 Gaps: 12

US-09-667-170A-440 (1-2239) x E89921 (1-440)

QY 325 TCAGAGAGAGAGAAAGTGCAGCTGAAGAGAAAGTGCACCTTACTGAGGGAGTCTCCATT 384
 Db 3 AsnGlyLysGlu-----LeuGlnLysAsnIleGlyPheSerAlaPheAlaIle 19
 QY 385 ATCAATGGCACCACATTTGGAGGAGGAATCTTCATCTCTCCCTAAAGGCGCTGCCAGAAC 444
 Db 20 ValMetGlyThrValIleGlySerGlyValPhePheLysIleSerAsnValThrGluVal 39
 QY 445 ACGGGCAGCGTGGCATGCTCTGACCATCTGCAGCGTGTGGGGTCTGTCACATATT 504
 Db 40 ThrGlyThrAlaGlyMetAlaLeuPheValTrpPheLeuGlyGlyIleIleThrIleCys 59
 QY 505 GGAGCTTTGCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCATTACACA 564
 Db 60 AlaGlyLeuThrAlaAlaGluLeuAlaAlaIleProGluThrGlyGlyLeuThrLys 79
 QY 565 TATATTTTGAAGCTTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCCTC 624
 Db 80 TyrIleGluThrThrGlyAspPheThrGlyPheLeuSerGlyTrpAlaGlnSerPhe 99
 QY 625 ATAATACGCGCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGTACATCTCGAA 684
 Db 100 IleTyrPheProAlaAsnValAlaAlaLeuSerIleValPheAlaThrGlnLeuIleAsn 119
 QY 685 CCATTTTATTCATGAGTAA-----ATCCCTGAACCTTGGCATCAAGCTCATTACAGCT 738
 Db 120 LeuPheHisLeuSerIleGlySerLeuIlePro-----IleAlaIleAlaSerAla 136
 QY 739 GTGGGCATAACTGTAGTGATGCTCTAAATACCATGAGTGTGACGTGGAGCGCCGATC 798
 Db 137 LeuSerIleValLeuIleAsnPheLeuGlySerLysAla-----GlyGlyIleLeu 153
 QY 799 CAGATTTTCTTAACCTTTTGCAGCTCACAGCAATCTGATAATTATAGTCCCTGGAGTT 858
 Db 154 GlnSerValThrLeuValIleLysLeuIleProIleIleValIleValIlePheGlyIle 173
 QY 859 ATGCAG-----CTAATTAAGGTCAAAGCAGAGAACTTTAAAGAC 897
 Db 174 PheGlnSerGlyAspIleThrPheSerLeuIlePro---ThrThrGlyAsnSerGlyAsn 192
 QY 898 GCCTTTTTCAGGAAGAGATTCAAGTATTACGCGGTTCGCACTGGCTTTTATTATGGAATG 957
 Db 193 GlyPhe-----PheThrAlaIleGlySerGlyLeuLeuAlaThrMet 206
 QY 958 TATGATATGCTGGTGTGTTTACTCAACTTTCTTACTGAAGAAGTAGAAACCTCGAA 1017
 Db 207 PheAlaTyrAspGlyTrpIleHisValGlyAsnValAlaGlyGluLeuLysAsnProLys 226
 QY 1018 AAAACCATTCCTCTGCAATATGATATCCATGGCCCATTTGTCCACCATTTGGCTATGCTG 1077
 Db 227 ArgAspLeuProLeuAlaIleSerValGlyIleGlyCysIleMetAlaValThrLeuLeu 246
 QY 1078 ACAATGTGGCCTACTTTACGACCATTAATGCTGAGAGAGTGTGTTTCAATGCAAGTG 1137
 Db 247 IleAsnAlaThrPheLeuLeuThrLeuProIle---GluLeuLeuAlaGlyAsnLeuAsn 265
 QY 1138 CGAGTGACCTTTTCTGAGCGGCTACTG-----GGAAATTTCTCATACAGTCTCCGATC 1191
 Db 266 AlaAlaSerAspThrSerLysIleLeuPheGlyGluAsnGlyGlyLysIleIleThrIle 285
 QY 1192 TTTGTCCTCTCTCTGCTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCTCCAGGTGA 1251
 Db 1192 TTTGTCCTCTCTCTGCTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCTCCAGGTGA 1251

Db 286 GlyIleLeuIleSerValTyrGlyThrIleAsnGlyTyrThrMetThrGlyMetArgVal 305
 QY 1252 TTCTATGTGGTCTCGAGAGGTCTACCTTTCCA-----GAAATCTCTCTCCATGATTCAT 1305
 Db 306 ProTyrAlaMetAlaGluArgLysLeuLeuProPheSerHisLeuPheAlaLysLeu--- 324
 QY 1306 GTCCGCAAGCACCTCTCTACAGCTGTTATTGTTTGGTTCACCTTTGACCAATGATAATG 1365
 Db 325 ThrLysSerGlyAlaProTrpPheGlyAlaIleIleGlnLeuIleIleAlaIleMet 344
 QY 1366 CTCTTCTCTGGACCTCGACAGCTCTTTGAATTTCTCAGTTTCTCCAGGTGGCTTTT 1425
 Db 345 MetSerMetGlyAlaPheAspThrIleThrAsnMetLeuIlePheValIleTrpLeuPhe 364
 QY 1426 ATTGGGCTGGCAGTGTGGCTGATTTATCTTCGTACATAAATGCCAGATATGCATCT 1485
 Db 365 TyrCysMetSerPheValAlaValIleIleLeuArgLysArgGluProAsnMetGluArg 384
 QY 1486 CCTTTCAAGTGGCAGCTGTTCT-----ATCCAGCTTTTCTTCCATGCCTCTTC 1539
 Db 385 ProTyrLysValProLeuTyrProIleIleProLeuIleAlaIleLeuAlaGlySerPhe 404
 QY 1540 ATGGTTGCCCTTCCCTCTATTGGACCATTTAGTACAGGATTTGGTTCGTCTCATCACT 1599
 Db 405 ValLeuIleAsnThrLeuPheThrGlnPheIleLeuAlaIleGlyIleLeuIleThr 424
 QY 1600 CTGACTGGAGTCCCTCGTATTAT 1623
 Db 425 AlaLeuGlyLeuProValTyrTyr 432

RESULT 12

AII1995

amino acid transporter alr1519 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AII1995

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AII1995

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-455 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA077885.1; PID:gl17135339; GSPDB:GN00179
 A:Experimental source: strain PCC 7120

A:Genetics:

A:Gene: alr1519

Alignment Scores:

Pred. No.: 5.78e-33 Length: 455
 Score: 463.00 Matches: 120
 Percent Similarity: 47.28% Conservative: 97
 Best Local Similarity: 26.14% Mismatches: 208
 Query Match: 11.24% Indels: 34
 DB: 2 Gaps: 9

US-09-667-170A-440 (1-2239) x AII1995 (1-455)

QY 322 TTTTCAGAGAGAGAAAGTGCAGCTG-----AAGAGAAAGTCACT 363
 Db 7 TyrSerLeuGluGlnIleGlnThrThrGlyAlaAspAlaProLysProLeuThr 26
 QY 364 TTACTAGGGAGTCTCCATTATTCATGGCACCATTCATGGAGCAGCAATCTTCATCTCT 423
 Db 27 LeuSerAspAlaValAlaLeuValGlyIleValIleGlyValGlyIlePheGlnThr 46
 QY 424 CCTAAGGCGCTGCTCCAGAACACAGCGGCGGCGATGTCTGTACCATCTGGAGCGTG 483
 Db 47 ProAlaLeuValAlaSerGlnAlaGlySerAspThrAlaValLeuLeuPheTrpLeuAla 66

Qy	484	TGTGGGGTCTCTCACTATTTTGGAGCTTGTCTTATGCTGAATTTGGGAACAACACTATAAAG	543
Db	67	GlyGlyIleValSerIleIleGlyAlaLeuCysTyAlaGluLeuAlaThrThrTyrPro	86
Qy	544	AAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGCTCCATTACCAAGCTTTTGTA	603
Db	87	AsnValGlyGlyAlaTyrTyrTyrLeuLysArgAlaPheGlyGlnAsnThrAlaPheLeu	106
Qy	604	CGAGTCTGGTGGCAACTCCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCA	663
Db	107	PheAlaTrpAlaArgLeuThrValIleGlnThrGlySerIleAlaLeuAlaAlaPheVal	126
Qy	664	TTTGGAGCGTCATCTCTGGACCACTTTTATTCAATGTCAATCCCTGAATCGGCATC	723
Db	127	PheGlyAspTyrAlaSerGluIleTrp-----ArgLeuGlyThrPheSerSer	142
Qy	724	AAGCTCATTAACAGCTGTGGGCATACTAGTGTAGTGTCTTAATAGCATGAGTCTCAGC	783
Db	143	SerMetTyrAlaAlaValIleIleAlaLeuLeuThrIleLeuAsnIleLeuGlyLeuHis	162
Qy	784	TGGAGCGCCGGATCCAGTATTTCTTACCTTTTTCACAGCTCAGCAATCTCGATAAT	843
Db	163	GlnGlyLysTrpThrGlnAsnLeuLeuThrAlaAlaGlnValLeuGlyLeuLeuVal	182
Qy	844	ATAGTCCCTCGAGTATTATGCAGCTAATTAAGGCTCAACGCGAGCAACTTTAAAGACGCTTT	903
Db	183	ValLeuPheGly-----LeuAlaSerThrAlaAsnSerAlaAsnSerAlaVal	198
Qy	904	TCAGAAAGAGATTCAAGTATTACGGCGTTGCCACTGGCTTTTATTATGAATGATGCA	963
Db	199	SerProGluProSerSerGlySerTrpGlyLeuAlaMetValPheValLeuLeuSer	218
Qy	964	TATCCTCGCTGGTGTTCACCTCAACTTTGTACTGAAGAAGTAGAAAAACCTGAAAAAAC	1023
Db	219	TyrGlyGlyTrpAsnGluAlaLayrIleSerAlaGluIleLysAsnArgGlnArgAsn	238
Qy	1024	ATTCCTCTGCAATATGTATCCATGGCCATTGTCCACCATGCTGCTGTGCAAAAT	1083
Db	239	IleLeuArgSerLeuMetTrpSerIleGlyIleIleThrAlaIleTyrLeuLeuIleAsn	258
Qy	1084	GTGGCTACTTTACGACCATTAACTGAGGAGCTGCTCTTCAAAATGCAGTGGCAGTG	1143
Db	259	LeuAlaPheLeuArgGlyLeuAlaAsnMetAlaAsnSerSerAlaValAlaAla	278
Qy	1144	ACCTTTTCTGAGCGCTACTGGAAATTTCTCATTAGCA---GTTCCGATCTTTGTGCC	1200
Db	279	AspLeuMetArgAlaValTrpGlyThrProGlyValValPheIleSerLeuLeuIleAla	298
Qy	1201	CTCTCCTCTTTGCTCCATGACGCTGTGTGTGTCTGTCTCCAGCTATTCTATGTT	1260
Db	299	IleCysAlaLeuGlyThrIleAsnAlaSerIlePheThrGlyAlaArgThrAsnTyrAla	318
Qy	1261	GGGTCTCGAGAGGTCACTTCACGAAATCTCTCCATGATTCATGTCGCGCAACACACT	1320
Db	319	LeuGlyGlnAspPheAsnLeuPheGlyPheMetGlySerTrpArgGlnIleProSerThr	338
Qy	1321	CCTCTACCAAGCTGTATTGTTTTGCACCCCTTTTACACATCATATGCTCTCTCTGGA---	1377
Db	339	ProAlaThrAlaLeuLeuValGlnGlyAlaIleAlaLeuAlaLeuValValLeuGlyThr	358
Qy	1378	-----GACCTCGCAGCTCTTTTGAATTTTCTCAGTTTTCGACGTTGGCTTTT	1425
Db	359	PheThrArgLysGlyPheGluThrMetValAspTyrThrAlaProValPheTrpPhePhe	378
Qy	1426	ATTGGCTGGCAGTGTCTGGCTGATTTATCTTCGATACAAATGCCAGATATCATCGT	1485
Db	379	PheLeuLeuSerGlyIleSerLeuLeuIleLeuArgGlnGlyGluProHisIleProArg	398
Qy	1486	CTTTTCAAGTCCCACTGTC-----ATCCCCAGCTTTTGTTCCTTCATGCTCCTTC	1539
Db	399	ProPheArgValIleProPheTyrProIleThrProLeuPhePheCysAlaValCysGlyTyr	418
Qy	1540	ATGGTTTGCCCTTTCCTCTATTTCGGACCCCATTTAGTACAGGGATTGGCTTCTC	1593

[illegible]

Db 153 ThrIleLeuLysValAlaProValLeuPheAlaValLeuGlyAlaIleHisLeuAla 172
 QY 871 AAAGTGCAA-----ACCAGCACTTTAAAGACGCCCTTTCA-----GGAAGA 912
 Db 173 LeuAsnProGlyLeuLeuValSerAsnYThrProAlaAlaProMetGlyLeuGlyAla 192
 QY 913 GATTCAGTATTACGGGTGGCCACTGGCTTTTATTATGGAATGTATGCATATGCTGC 972
 Db 193 LeuGlyThrValThrValLeuVal-----PheTrpAlaYrValGly 206
 QY 973 TGGTTTACCTCACTTTGTACTGAAGAAGTAGAAGAACCTGAAACCAATCCCTTT 1032
 Db 207 PheGluLeuValThrValProAlaAspGluValArgAspProGluArgThrIleProLeu 226
 QY 1033 GCAATATGTATATCCATGGCCATGTCCACCATGGCTATGTGCTGACAAATGGCCCTAC 1092
 Db 227 SerIleThrLeuGlyMetIlePheValThrPheYrIleLeuThrAsnAlaValIle 246
 QY 1093 TTTAGGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCAGCTTTTCT 1152
 Db 247 LeuGlyLeuValProTrpArgValLeuAlaSerSerThrAlaProLeuThrValAlaGly 266
 QY 1153 GAGCGCTACTGGGAATTTTC---TCATTAGCAGTTCCCATCTTTGTCCTCTCTCCG 1209
 Db 267 TyrSerLeuMetGlyGlyIleGlyAlaLeuIleLeuThrAlaGlyAlaValPheSerIle 286
 QY 1210 TTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTATTCTATGCTTCCGCTCGA 1269
 Db 287 AlaGlySerGluGluAlaGlyMetLeuThrThrAlaArgLeuLeuPheAlaMetSerGlu 306
 QY 1270 GAGGTCACCTCCGAATCTCTCCATGATTCATGTCGCGAAGCAGCTCTCTACCA 1329
 Db 307 AspGlyPheLeuProGlyPheLeuSerArgValHisArgPheGlyThrProHisMet 326
 QY 1330 GCTGTTATTGTTTTCACCTTTGACAATGATAATGCTCTCTCGGAGACCTCGACAGT 1389
 Db 327 SerIleLeuValGlnAsnLeuThrAlaLeuLeuAlaLeuThrGlyThrValSerGly 346
 QY 1390 CTTTTCGAATTCCTCAGTTTGGCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTG 1449
 Db 347 LeuIleGluLeuSerValValThrLeuLeuLeuProTyAlaValThrCysIleSerLeu 366
 QY 1450 ATTTATCTTCGATACAAATGCCAGATATGATGCTCTCTTCAGAGTGGCCACTGTCATC 1509
 Db 367 AlaIleLeuArgArg---AspGlySerGlyIleProLeuLysSerValLeuGlyVal 385
 QY 1510 CCAGCTTTGTTTTCCTTCACATSCCTCTCATGTTGGCTTCCCTCTATTCGAGCCCA 1569
 Db 386 -----LeuValCysIleYrLeuLeu-----MetAsnThrPro 397
 QY 1570 TTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTGCTATTAFTCTTT 1629
 Db 398 SerThrAlaTrpGlyLeuLeuLeuIleLeuSerGlyAlaProLeuThrLeuIlePhe 417

RESULT 14
 S61943
 methionine transport protein, high affinity - yeast (Saccharomyces cerevisiae)
 N:Alternate names: methionine permease; protein G4340; protein YGR055W
 C:Species: Saccharomyces cerevisiae
 C>Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jun-2000
 C:Accession: S61943; S64349
 R:Isnard, A.D.; Thomas, D.; Surdin-Kerjan, Y.
 submitted to the EMBL Data Library, November 1995
 A:Reference number: S61943
 A:Accession: S61943
 A:Molecule type: DNA
 A:Residues: 1-574 <ISN>
 A:Cross-references: EMBL:U40316; NID:g1101906; PID:AAB63529.1; PID:g1101907
 A:Experimental source: strain X2180-1A
 R:Entian, K.D.; Rose, M.; Koetter, P.; Roelmer, A.; Sehrsam, I.; Hempel, S.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64335
 A:Accession: S64349

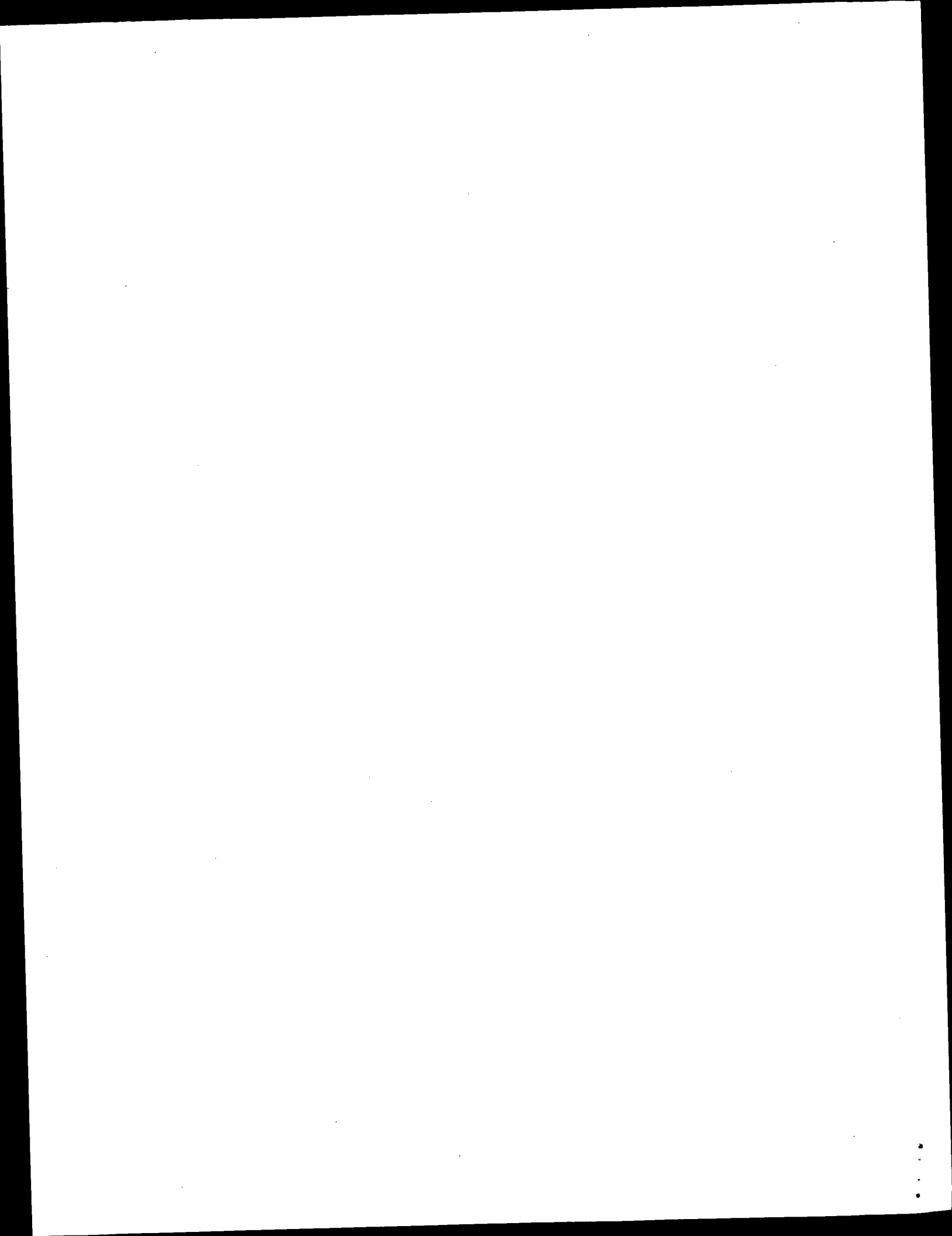
A:Molecule type: DNA
 A:Residues: 1-574 <ENT>
 A:Cross-references: EMBL:Z72840; NID:g1323066; PIDN:CAA97055.1; PID:g1323067; MIPS:YG
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:MUP1
 A:Cross-references: SGD:S0003287; MIPS:YGR055W
 A:Map position: 7R
 C:Superfamily: hypothetical protein YHL036W
 C:Keywords: amino acid transport; transmembrane protein
 F:92-108/Domain: transmembrane #status predicted <TM1>
 F:140-156/Domain: transmembrane #status predicted <TM2>
 F:183-199/Domain: transmembrane #status predicted <TM3>
 F:208-224/Domain: transmembrane #status predicted <TM4>
 F:296-312/Domain: transmembrane #status predicted <TM5>
 F:318-334/Domain: transmembrane #status predicted <TM6>
 F:343-359/Domain: transmembrane #status predicted <TM7>
 F:390-406/Domain: transmembrane #status predicted <TM8>
 F:457-473/Domain: transmembrane #status predicted <TM9>
 F:494-510/Domain: transmembrane #status predicted <TM10>

Alignment Scores:
 Pred. No.: 1,75e-24 Length: 574
 Score: 368.50 Matches: 135
 Percent Similarity: 45.61% Conservative: 104
 Best Local Similarity: 25.76% Mismatches: 199
 Query Match: 8.94% Indels: 86
 DB: 2 Gaps: 22

US-09-667-170a-440 (1-2239) x S61943 (1-574)
 QY 280 GAGCCCTTTTCAGGAAGACGCCCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAA 339
 Db 36 AspAlaAspAsnGlyAlaSerAspPheGluAlaGlyGlnPheAla---ThrGluLeu 54
 QY 340 GTCCACTCAAGAGGAAGTCACTTACTGAGGGGAGTCTCCATTATCATTTGGCACCATC 399
 Db 55 AspGlnGlyGluGlnLeuGlyIleLeuSerCysIleGlyLeuIleCysAsnArgMet 74
 QY 400 ATTGGAGCAGGAATCTTCATCTCTCCTAAGGCGTGTCTCCAGAACACACGCGGTGGGC 459
 Db 75 LeuGlyThrGlyValPheAlaValSerSerThrIleYrThrLeuCysGlySerValGly 94
 QY 460 ATGTCTCTCACCATCTGGACGCTGTGGGTCTGTCCTGTCACATTTGGAGCTTTCTCTAT 519
 Db 95 LeuAlaLeuIleMetTrpAlaValGlyAlaIleAlaIleSerGlyLeuYrValYr 114
 QY 520 GCTGAATTGGGAACAACATAAAGAAATCTGGAGTCTATTACACATATATTTTGGAGTC 579
 Db 115 MetGluPheGlyThrAlaIleProLysAsnGlyGlyGluLysAsnYrLeuGluAlaIle 134
 QY 580 TTTGGTCCATTACACAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGCCCTGCA 639
 Db 135 PheArgLys---ProLysPhe----- 140
 QY 640 GCTACTGCTGTGATACCTCGCATTTGGAGCGCTACATCTTGGAACCATTTTATTATCAA 699
 Db 141 -----PheIleThrCysMetYrAlaAlaYrIle-----PhePheLeuGly 154
 QY 700 TGTGAATCCCT-----GAACCTTGGCATCAAGCTCATTCAGCT----- 738
 Db 155 TrpAlaAlaGlyAsnSerIleAsnThrAlaIleMetPheLeuThrAlaAlaAspThrGlu 174
 QY 739 -----GTGGGCATACTACTAGTG-----ATGGTCCCTA 765
 Db 175 ValThrLysTrpAsnGlnArgGlyIleGlyValAlaValPhePheAlaPheLeuIle 194
 QY 766 AATAGCATGAGTGTGACGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTCAAGCTC 825
 Db 195 AsnSerLeuAsnValGlyIleGlyLeuYrLeuGlnAsnIleLeuGlyIlePheYsIle 214
 QY 826 ACAGCAATTCGATTAATATAGTCCCTGGAGTTATGCGAGCTAATAAAGGT----- 876

Db 248 sPropelleProGluProThrGlyGln-----ProGlyGluPh 261
QY 945 TTATTATGGA-----ATGATGTCATATGCTGGTGT 977
Db 261 eGlyIleGlyIlePheArgGlyAlaAlaIleIlePheAlaTyrValGlyPheG1 281
QY 978 TTACCTCAACTTTGTTACTGAAGAAGTAAACCCCTGAAACCAACCATTCCTCCCTGCAAT 1037
Db 281 uAlaValSerThrAlaAlaAlaGluAlaLysAsnProSerArgAspValProIleGlyI1 301
QY 1038 ATGTATATCCATGCCATTGTCACCATTTGGCTAT-----GTGCTGACAAA 1082
Db 301 eLeuGlyAlaLeuIleIleCysThrLeuIleTyrMetAlaValAlaAlaValMetThrG1 321
QY 1083 TGTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGTTTCAAAATGCGAGTGGCAGT 1142
Db 321 yValValProPhe-----ArgGluLeuAlaSerProAlaProIleAlaVa 336
QY 1143 GACCTTTTCTGACGGCTA----- 1161
Db 336 lAlaIle---AspArgMetGlyLeuGluTrpAlaAspIleProTyrAlaAlaGluG1 355
QY 1162 -----CTGGGAATTTCTCATTTAGCAGTCCGATCTTTGCTGGCTCTCC-- 1206
Db 355 yGlyLysLeuAsnLeuSerPheAlaIleLysIleGlyAlaIleThrGlyLeuSerSe 375
QY 1207 -----TGTGTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCTCCAGGTT 1250
Db 375 rValMetLeuValLeuCysTyrGlyGln-----ThrArgI1 387
QY 1251 ATTCTATGTTGGTCTCGAGAGGTCACCTTCAGAAATCCTCTCCATGATTCANGTCGG 1310
Db 387 ePheTyrThrMetAlaArgAspGlyLeuLeuProLysValPheAlaGluIleHisProLy 407
QY 1311 CAAGCACACTCCTCTACACGCTGTTATGTTTGCAC----- 1347
Db 407 sPheArgThrProTrpLeuGlyThrIleLeuLeuGlyValValIleAlaIleAlaAlaSe 427
QY 1348 -----CCTTTGACAATGATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTT 1400
Db 427 rPheLeuProIleSerLeuLeu-----GlyAspLeuValSerLeuGlyThrAl 443
QY 1401 CCTCAGTTTGGCCAGTGGCTTTTATGTTGGCTGGCAGTTGCTGGGCTGATTATCTTCG 1460
Db 443 aValAlaPheSer-----IleValCysLeuSerValIleTyrLeuAr 457
QY 1461 ATCAAAATGCCAGATATGATCCTCTTTCAAGTGCCA-----CTGTTTCATCCCGC 1514
Db 457 gIleLysHisProAspLeuProArgProPheLysValProGlyIlePheThrAlaAl 477
QY 1515 TTTGTTTTCCTTCACATGCTCTTC-----ATGTTTGGCCT 1550
Db 477 aAlaGlyIleAlaAlaCysLeuPheLeuProTyrGlnAsnPheGlnProMetIleValHi 497
QY 1551 TTCCTCTATTCGGACCCATTTAGTACAGGATGGCTTCGTCATCACTCTGACTGGAGT 1610
Db 497 sAlaMetAsnAspAsnProLeuProLeuMetIleLeuGlyClyTyrAlaAlaValGlyAl 517
QY 1611 CCCTGCGTATTATCTCTTTATATATGGACAGAAA 1647
Db 517 alleIleTyrIleAlaTyrGlyTyrTrpHisSerLys 529

Search completed: April 16, 2003, 16:26:07
Job time : 65 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 14:34:15 ; Search time 25 Seconds
(without alignments)
7429.240 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+n2p.model -DEV-xmlh
-Q/cgn21/USPTO_spool/US096671170/runat_04042003_090917_20285/app_query.fasta_1.2375
-DB-SwissProt_40 -QFMT-fastan -SUFFIX-n2p.rsp -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS-bits -SPART=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=PTO -NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US096671170.ecgn_1_24_@runat_04042003_090917_20285 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7
-YGAPOP=10 -YGAPEXT=0 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match	Length	DB ID	Description				
1	2519	61.1	501	1	XCT_HUMAN	Q9up55 homo sapien			
2	2307	56.0	501	1	XCT_MOUSE	Q9wtr6 mus musculus			
3	1210.5	29.4	507	1	LAT1_HUMAN	Q01650 homo sapien			
4	1204	29.2	512	1	LAT1_RAT	Q01650 rat mus musculus			
5	1202	29.2	512	1	LAT1_MOUSE	Q92127 mus musculus			
6	1131.5	27.5	511	1	YLA1_HUMAN	Q9um01 homo sapien			
7	1108	26.9	533	1	LAT2_RAT	Q9wvr6 rat mus musculus			
8	1103	26.8	531	1	LAT2_MOUSE	Q9qxr9 mus musculus			
9	1101.5	26.7	535	1	LAT2_HUMAN	Q9uh15 homo sapien			
10	1075.5	26.1	487	1	BAT1_HUMAN	P82251 homo sapien			
11	1074.5	26.1	487	1	BAT1_MOUSE	P82252 rat mus musculus			
12	1067.5	25.9	487	1	BAT1_RAT	Q9qxa6 mus musculus			
13	368.5	8.9	574	1	MUP1_MOUSE	P50276 saccharomyc			
14	343.5	8.3	546	1	MUP3_YEAST	P38734 saccharomyc			
15	321	7.8	445	1	YHFM_ECOLI	P45539 escherichia			
16	300.5	7.3	440	1	YJ99_ECOLI	Q10858 mycobacteri			
17	293.5	7.1	658	1	CTRL_HUMAN	P52569 homo sapien			
18	286	6.9	629	1	CTRL_MOUSE	P30825 homo sapien			

19	279.5	6.8	622	1	CTRL_MOUSE	Q09143 mus musculus
20	275.5	6.7	624	1	CTRL_RAT	P30823 rattus norv
21	273	6.6	657	1	CTRL_MOUSE	P18581 mus musculus
22	272.5	6.6	430	1	YBAT_ECOLI	P77400 escherichia
23	253	6.1	636	1	CTR4_HUMAN	O43246 homo sapien
24	252.5	6.1	459	1	AAPG_BACSU	O06005 bacillus su
25	251.5	6.1	458	1	YDGF_BACSU	P67074 bacillus su
26	250.5	6.1	497	1	ANSP_SALTY	P40812 salmonella
27	239	5.8	489	1	ANSP_MCTU	O33261 mycobacteri
28	236.5	5.7	499	1	ANSP_ECOLI	P77610 escherichia
29	232	5.6	457	1	PROX_ECOLI	P77327 escherichia
30	232	5.6	467	1	MMUP_ECOLI	P18696 emericella
31	232	5.5	550	1	PUTX_EMENI	O58026 methanococc
32	228.5	5.5	435	1	Y609_METJA	P37460 salmonella
33	227.5	5.5	456	1	PROX_SALTY	O10875 mycobacteri
34	224.5	5.4	481	1	YJ79_MCTU	P24170 escherichia
35	222.5	5.4	439	1	POTE_ECOLI	P15993 escherichia
36	222.5	5.4	457	1	AROP_ECOLI	O60170 schizosacch
37	222.5	5.4	574	1	ME22_SCHPO	P34054 trichoderma
38	222	5.4	573	1	INAI_TRIHA	P44768 haemophilus
39	221	5.4	435	1	POTE_HAEIN	O87394 rhizobium m
40	220	5.3	465	1	Y093_RHIME	P24877 saccharomyc
41	218	5.3	611	1	LYPL_YEAST	P25737 escherichia
42	216.5	5.3	488	1	LYSP_ECOLI	P23891 escherichia
43	214	5.2	444	1	CADB_ECOLI	P39269 escherichia
44	214	5.2	445	1	YUDE_ECOLI	P48813 saccharomyc
45	214	5.2	663	1	GNPL_YEAST	

ALIGNMENTS

RESULT 1									
XCT_HUMAN					STANDARD; PRT; 501 AA.				
ID	Q9UPY5;								
DT	16-OCT-2001 (Rel. 40, Created)								
DT	16-OCT-2001 (Rel. 40, Last sequence update)								
DT	15-JUN-2002 (Rel. 41, Last annotation update)								
DE	Cystine/glutamate transporter (Amino acid transport system xc-)								
DE	(Calcium channel blocker resistance protein CCBRI).								
GN	SLC7A11.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
OX	NCBI_TaxID=9606;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	MEDLINE=21082210; PubMed=11213471;								
RA	Sato H., Tamba M., Kuriyama-Matsumura K., Okuno S., Bannai S.;								
RT	"Molecular cloning and expression of human xCT, the light chain of								
RT	amino acid transport system xc-";								
RL	Antioxid. Redox. Signal. 2:665-671(2000).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RA	Conklin D.S., Beach D.H.;								
RT	"CCBR1, novel CD98 light chain implicated in redox control and calcium								
RT	signaling.";								
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.								
RN	[3]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Placenta;								
RA	Chaney C.D., Kekuda R., Wang H., Huang W., Prasad P.D., Smith S.B.,								
RA	Ganapathy V.;								
RT	"Structure, function and regulation of human cystine/glutamate								
RT	transporter in retinal pigment epithelial cells.";								
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.								
RN	[4]								
RP	SEQUENCE FROM N.A.								
RA	Borsani G., Manzoni M., Palacin M., Pineda M., Gasol E.;								
RA	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.								
RN	[5]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Kidney;								

RA Strausberg R.; to the EMBL/GenBank/DBJ databases.
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY EXCHANGE OF ANIONIC
 CC AMINO ACIDS WITH HIGH SPECIFICITY FOR ANIONIC FORM OF CYSTINE AND
 CC GLUTAMATE (BY SIMILARITY).
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
 CC TRANSPORT PROTEIN SLC3A2/4F2HC (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
 CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
 CC SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; AB026891; BAA82628.1; -
 CC EMBL; AF200708; AAG35592.1; -
 CC EMBL; AF252872; AAK49111.1; -
 CC EMBL; AJ277882; CAC81905.1; -
 CC EMBL; BC012087; AAH12087.1; -
 CC Genew; HGNC:11059; SLC7A11.
 CC InterPro; IPR002293; AA/rel.pmeasel.
 CC InterPro; IPR004760; LAA.transprot.
 CC InterPro; IPR004841; Permease.
 CC Pfam; PF00324; aa_permeases; 1.
 CC TIGRFAMs; TIGR00911; 2A0308; 1.
 CC Transprot; Amino-acid transport; Transmembrane.
 CC TRANSMEM 44 64 POTENTIAL.
 CC FT TRANSMEM 76 96 POTENTIAL.
 CC FT TRANSMEM 136 156 POTENTIAL.
 CC FT TRANSMEM 159 179 POTENTIAL.
 CC FT TRANSMEM 190 210 POTENTIAL.
 CC FT TRANSMEM 235 255 POTENTIAL.
 CC FT TRANSMEM 266 286 POTENTIAL.
 CC FT TRANSMEM 318 338 POTENTIAL.
 CC FT TRANSMEM 365 385 POTENTIAL.
 CC FT TRANSMEM 388 408 POTENTIAL.
 CC FT TRANSMEM 423 443 POTENTIAL.
 CC FT TRANSMEM 450 470 POTENTIAL.
 CC FT CARBOHYD 314 314 N-LINKED (GLCNAC...) (POTENTIAL).
 CC SQ SEQUENCE 501 AA; 55422 MW; 3EF2648B9A9F59E CRC64;

Alignment Scores:
 Pred. No.: 2,22e-190 Length: 501
 Score: 2519.00 Matches: 498
 Percent Similarity: 96.51% Conservative: 0
 Best Local Similarity: 96.51% Mismatches: 0
 Query Match: 61.14% Indels: 18
 DB: 1 Gaps: 1

US-09-667-170a-440 (1-2239) x XCT_HUMAN (1-501)

QY 184 AAGCTGTTGTCACCACTCTCCAAAGGAGGTTCACCTGACGGAAATGTTAACGGAGG 243
 Db 4 LysProValValSerThrIleSerLysGlyGlyTyrLeuGlnGlyAsnValAsnGlyArg 23
 QY 244 CTGCTCTTCCCTGGGCAACAGGAGGACCCCTGGGAGGAGCCCTTTTCAGGAAGAGACGCC 303
 Db 24 LeuProSerLeuGlyAsnLysGluProProGlyGln- 35
 QY 304 TTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
 Db 36 -----GluLysValGlnLeuLysArgLysValThr 45
 QY 364 TTAAGTGGGGAGGAGTCTCCATTATCATGGACCATCATGGAGGAGGAGATCTTCATCTCT 423
 Db 46 LeuLeuArgGlyValSerIleIleIleGlyThrIleIleGlyAlaGlyIlePheIleSer 65

QY 424 CCTAAGGGCGTGTCTCCAGAACACGGCGGCGTGGCATGTCTCTACCATCTGACGCGTG 483
 Db 66 ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleTrpThrVal 85
 QY 484 TGTGGGGTCTGTCTACTATTGAGCTTTGCTTATGCTGAATGGGGAACAACATAAG 543
 Db 86 CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrThrIleLys 105
 QY 544 AATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTACCAAGCTTTTGT 603
 Db 106 LysSerGlyGlyHisTyrThrIleLeuGluValPheGlyProLeuProAlaPheVal 125
 QY 604 CGAGTCTGGGTGGAACCTCTCATATACGCCCTCGAGCTACTGTGTGATATCCCTGGCA 663
 Db 126 ArgValTrpValGluLeuLeuIleIleArgProAlaAlaThrAlaValIleSerLeuAla 145
 QY 664 TTTGGAGCGTACATCTCGGAACCATTTTATTCATGTGAATCCCTGAACTCCGATC 723
 Db 146 PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle 165
 QY 724 AAGCTCATTACAGCTGTGGGCATACTGTAGTGTCTTAAATAGCATGAGTGTCTCAGC 783
 Db 166 LysLeuIleThrAlaValGlyIleThrValValMetValLeuAsnSerMetSerValSer 185
 QY 784 TGGAGCGCCCGGATCCAGATTTTCTTAACTTTTGAAGCTCACACCAATTTCTGATATT 843
 Db 186 TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIleIle 205
 QY 844 ATAGTCCCTGGAGTTATGACAGTAAATTAAGGTCAACAGCAGAACCTTTAAAGACGCTTT 903
 Db 206 IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPheLysAspAlaPhe 225
 QY 904 TCAGGAGAGATTCAGATTTATACCGGTTGCCACTGGCTTTTATTATGATGATGTATGCA 963
 Db 226 SerGlyArgAspSerSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla 245
 QY 964 TATGCTGCTGGTGTTCACCTCACTTTTGTACTGAAAGCTAGAAACCCCTGAAAAACC 1023
 Db 246 TyrAlaGlyTrpPheTyrLeuAsnPheValThrGluGluValGluAsnProGluLysThr 265
 QY 1024 ATTCCTCTTGAATATGATATATCCATGCCATGTCCACCATGTGCTATGTGTGACAAAT 1083
 Db 266 IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn 285
 QY 1084 GTGGGCTACTTTACAGACCAATTAATGCTGAGGAGCTGTCTTCAATGCAAGTGGCAGTG 1143
 Db 286 ValAlaTyrPheThrThrIleAsnAlaGluGluLeuLeuSerAsnAlaValAlaVal 305
 QY 1144 ACCTTTTCTGAGCGGCTACTGGGAAATTTCTATTACAGTTCCGATCTTTGTTGCCCTC 1203
 Db 306 ThrPheSerGluArgLeuLeuGlyAsnPheSerLeuAlaValProIlePheValAlaLeu 325
 QY 1204 TCCTGCTTTGGTCTCCATGAACGGTGGTGTGTTGCTGCTCTCCAGTATTCTATGTTGCG 1263
 Db 326 SerCysPheGlySerMetAsnGlyGlyValPheAlaValSerArgLeuPheTyrValAla 345
 QY 1264 TCTCGAGAGGTCACCTCCAGAAATCTCTCATGATGATCATGTCCGCAAGCACACCTCT 1323
 Db 346 SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValArgLysHisThrPro 365
 QY 1324 CTACCAAGCTGTATTGTTTTCACCCCTTTGACAAATGATGATGCTCTCTCTCGAGACCTC 1383
 Db 366 LeuProAlaValIleValLeuHisProLeuThrMetIleMetLeuPheSerGlyAspLeu 385
 QY 1384 GACAGCTTTTGAATTTCTCAGTTTGTGCGAGTGGCTTTTATTGGCTGGCAGTTGCT 1443
 Db 386 AspSerLeuLeuAsnPheLeuSerPheAlaIaTgTLeuPheIleGlyLeuAlaValAla 405
 QY 1444 GGCTGATTTATCTTCGATACAAATGCCAGATATGATGATGCTCTTCAAGTGCCACTG 1503
 Db 406 GlyLeuIleTyrLeuArgTyrLysCysProAspMetHisArgProPheLysValProLeu 425
 QY 1504 TTCAATCCAGCTTGTGTTTCTTCCCTTCACATGCCCTCTTCATGGTTGCCCTTCTATTTCG 1563

```

Db 426 PheIleProAlaLeuPheSerPheThrCysLeuPheMetValAlaLeuSerLeuTyrSer 445
QY 1564 GACCCATTACTACAGGATGGCTTCGTCATCCTGACTGAGTCCCTCGGTATTTAT 1623
Db 446 AspProPheSerThrGlyIleGlyPheValIleThrLeuThrGlyValProAlaTyrTyr 465
QY 1624 CTCCTTTATATATGAGCAAGAACCCAGGTGGTTTAGAATAATGTCAGAGAAATAACC 1683
Db 466 LeuPheIleLeuTAspLysLysProArgTrpPheArgIleMetSerGluLysIleThr 485
QY 1684 AGACATTACAATAATACCTGGAGTGTACACAGAAAGAGATAAGTTA 1731
Db 486 ArgThrLeuGlnIleLeuGluValValProGluGluAspLysLeu 501

RESULT 2
XCT_MOUSE
ID XCT_MOUSE STANDARD; PRT; 502 AA.
AC QSWTR6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cystine/glutamate transporter (Amino acid transport system xc-) (XCT).
GN SLC7A11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=9223452; PubMed=10206947;
RA Sato H., Tamba M., Ishii T., Bannai S.;
RT "Cloning and expression of a plasma membrane cystine/glutamate
RT exchange transporter composed of two distinct proteins.";
RL J. Biol. Chem. 274:11455-11458 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC SPRAIN-129/SvJ; TISSUE=Liver;
RA Sasaki H., Sato H., Bannai S.;
RT "Isolation and functional characterization of mouse cystine/glutamate
RT exchange transporter gene.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY EXCHANGE OF ANIONIC
CC AMINO ACIDS WITH HIGH SPECIFICITY FOR ANIONIC EXCHANGE OF CYSTINE AND
CC GLUTAMATE.
CC -! SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
CC TRANSPORT PROTEIN SLC3A2/4F2HC.
CC -! SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -! SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB022345; BAA77220.1;
DR EMBL; AB037661; BAA90522.1;
DR EMBL; AB037650; BAA90522.1; JOINED.
DR EMBL; AB037651; BAA90522.1; JOINED.
DR EMBL; AB037652; BAA90522.1; JOINED.
DR EMBL; AB037653; BAA90522.1; JOINED.
DR EMBL; AB037654; BAA90522.1; JOINED.
DR EMBL; AB037655; BAA90522.1; JOINED.
DR EMBL; AB037656; BAA90522.1; JOINED.
DR EMBL; AB037657; BAA90522.1; JOINED.
DR EMBL; AB037658; BAA90522.1; JOINED.
DR EMBL; AB037659; BAA90522.1; JOINED.
DR EMBL; AB037660; BAA90522.1; JOINED.

```

```

MGI:1347355; Slc7a11.
DR InterPro: IPR002293; AA/rel_primease1.
DR InterPro: IPR004760; LAA transport.
DR InterPro: IPR004841; Permease.
DR Pfam: PF00324; aa_permeases; 1.
DR TIGRFAMs: TIGR00911; 2A0308; 1.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
FT TRANSMEM 235 255 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT TRANSMEM 365 385 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 423 443 POTENTIAL.
FT TRANSMEM 450 470 POTENTIAL.
FT CARBOHYD 19 19 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 502 AA; 55455 MW; A4A185102D83992A CRC64;

Alignment Scores:
Pred. No.: 1,05e-173 Length: 502
Score: 2307.00 Matches: 444
Percent Similarity: 93.01% Conservative: 35
Best Local Similarity: 86.21% Mismatches: 18
Query Match: 56.00% Indels: 18
DB: 1 Gaps: 1

US-09-667-170a-440 (1-2239) x XCT_MOUSE (1-502)
QY 184 AAGCCTGTTGTGTCCACCATCTCCAAAGGAGGTTCACCTGCAGGAGAAATGTTAACGGGAGG 243
Db 4 LysProValValAlaThrIleSerLysGlyGlyTyrLeuGlnGlyAsnMetSerGlyArg 23
QY 244 CTGCCTTCCTCCGGCAACAGGAGGACCTGGCGAGGACGGCTTTTCAGGAAGAGACGCC 303
Db 24 LeuProSerMetGlyAspGlnGluProGlyGln----- 35
QY 304 TTTCAGGAAGAGAGAGCGCTTTTCAGGAAGAGAGAGTGCAGCTCAAGAGCAAGATCACT 363
Db 36 -----GluLysValValLeuLysLysLysIleThr 45
QY 364 TTACTGAGGGAGTCTCCATTATCATTCGCACCATTCATTGGAGGAGGAATCTTCATCTCT 423
Db 46 LeuLeuArgGlyValSerIleIleIleGlyThrValIleGlySerGlyIlePheIleSer 65
QY 424 CTAAGGGCGTCTCCAGAACACGCGCGCTGCGCATGCTCTCGACCATCTGGACGGTG 483
Db 66 ProLysGlyIleLeuGlnAsnThrGlySerValGlyMetSerLeuValPheTrpSerAla 85
QY 484 TGTGGGGTCTGTCTACTATTTCGAGCTTTGTCTTATGCTGAATTCGGAACAACATAAAG 543
Db 86 CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrSerIleLys 105
QY 544 AAATCTGAGGTCATTACACATATATTTTGGAGTCTTTTGGTCCATTACAGGCTTTTGTGA 603
Db 106 LysSerGlyGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 125
QY 604 CAGTCTCGGTGGAACTCCTCATATACGCCCTCGACCTACTGCTGTGATATCCCTGGCA 663
Db 126 ArgValTrpValGluLeuLeuValIleIleIleIleIleIleIleIleIleIleIleIle 145
QY 664 TTGGACGCTACATCTCTGGAACCACTTTTATTCAATGTGAAATCCCTGAACCTTGGATC 723
Db 146 PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle 165
QY 724 AAGCTCATTACAGCTGTGGGCATTAACGTGTAGTGATGTCCTAAATAGCATGATGTCAGC 783
Db 166 LysLeuValThrAlaValGlyIleThrValValMetValLeuAsnSerThrSerValSer 185

```

Thu Apr 17 09:48:26 2003

QY	784	TGGAGCCCGCGATCCAGATTTCTTAAACCTTTTTCGAAGCTCACAGCAATTCGTGATAATT	843
Db	186	trpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIleIle	205
QY	844	ATAGTCCCTGGAGTATGACGCTAAATAAAGGTCAAACGACGAACTTTAAAGACGCTTT	903
Db	206	IleValProGlyValIleGlnLeuIleLysGlyGlnThrHisHisPheLysAspAlaPhe	225
QY	904	TCAGGAAGAGATTCAAGTATTACGCGGTTCGCCACTGGCTTTTATTATGAAGTGTATGCA	963
Db	226	SerGlyArgAspThrSerLeuMetGlyLeuProLeuAlaPheTyrTyrGlyMetTyrAla	245
QY	964	TATGCTGGCTGGTTTACCTCAACTTCTTACTGAAAGAAGTAGAAAACCCCTGAAAAACC	1023
Db	246	TyrAlaGlyTrpPheTyrLeuAsnPheIleThrGluGluValAspAsnProGluLysThr	265
QY	1024	ATTCCCTTGGCAATGATGATATGCCAATGTGCACCATTCGCTATGCTGCTGACAAAT	1083
Db	266	IleProLeuAlaIleCysIleSerMetAlaIleIleThrValGlyTyrValLeuThrAsn	285
QY	1084	GTGGCTACTTTACGACCACTTAATGCTCAGGAGCTGCTGCTTTCAATGAGTGGCAGTG	1143
Db	286	ValAlaTyrPheThrIleSerAlaGluLeuLeuGlnSerSerAlaValAlaVal	305
QY	1144	ACCTTTTCTGAGCGGCTACTCGGAAATTTCTCATTAGCAGTTCGATCTTTGTTGCCCTC	1203
Db	306	ThrPheSerGluArgLeuLeuGlyLysPheSerLeuAlaValProIlePheValAlaLeu	325
QY	1204	TCCTGCTTTGGCTCCATGAACGGTGGTGTGCTGCTCCAGGTATTCTATGTTGCG	1263
Db	326	SerCysPheGlySerMetAsnGlyGlyValPheAlaValSerArgLeuPheTyrValAla	345
QY	1264	TCTCGAGAGGGTCACCTCCAGAAATCTCTCCATGATTCATGTCGCGAAGCACACTCCT	1323
Db	346	SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValHisLysHisThrPro	365
QY	1324	CTACCACTGTTATGCTTTGGCACCTTTGACAAATGATAATGCTCTCTCTGGAGACCTC	1383
Db	366	LeuProAlaValIleValLeuHisProLeuThrMetValMetLeuPheSerGlyAspLeu	385
QY	1384	GACAGTCTTTGAATTCCTCAGTTTTCGAGGTGGCTTTTATTGGGCTGGCAGTTGCT	1443
Db	386	TyrSerLeuLeuAsnPheLeuSerPheAlaArgTrpLeuPheMetGlyLeuAlaValAla	405
QY	1444	GGGCTGATTATCTCGATACAAATGCCAGATATGCATCGCTTCCTTCAAGGTGCCACTG	1503
Db	406	GlyLeuIleTyrLeuArgTyrLysArgProAspMetHisArgProPheLysValProLeu	425
QY	1504	TTTCATCCCACTTGTTCCTTCACATGCCCTCTTCATGTTGGTGGCTTCCTCTATTTCG	1563
Db	426	PheIleProAlaLeuPheSerPheThrCysLeuPheMetValValLeuSerLeuTyrSer	445
QY	1564	GACCAATTTAGTACAGGATTCGCTCGTCATCACTCTGACTGGAGTCCCTGCGTATTAT	1623
Db	446	AspProPheSerThrGlyValGlyPheLeuIleThrLeuThrGlyValProAlaTyrTyr	465
QY	1624	CTCTTTTATTATGGGACAGAAACCCAGGTGGTTTAGAATAATATCTCAGAGAAAATAACC	1683
Db	466	LeuPheIleValTrpAspLysLysProLysTrpPheArgArgLeuSerAspArgIleThr	485
QY	1684	AGACATTACAAATATCTGGAAGTTGTACCAAGAGAGATAAG	1728
Db	486	ArgThrLeuGlnIleIleLeuGluValValProGluAspSerLys	500
RESULT 3			
LATI_HUMAN			
ID	LATI_HUMAN	STANDARD;	PRT; 507 AA.
AC	Q01650; Q9UBN8; Q9UP15; Q9UQC0;		
DT	01-JUN-1994 (Rel. 29, Created).		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Large neutral amino acids transporter small subunit 1 (L-type amino acid transporter 1) (4F2 light chain) (4F2LC) (CD98 light		

DE	chain) (Integral membrane protein El6) (hLAT1).	
GN	SLC7A5 OR LAT1 OR WPE16 OR CD98LC.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=98421678; PubMed=9751058;	
RA	Mastrobardino L., Spindler B., Pfeiffer R., Skelly P.J., Loffing J.,	
RA	Shoemaker C.B., Verrey F.;	
RT	"Amino-acid transport by heterodimers of 4F2hc/CD98 and members of a	
RL	permease family.";	
RL	Nature 395:288-291(1998).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Ovary;	
RA	Yanagida O., Segawa H., Miyamoto K., Takeda E., Goya T., Endou H.,	
RA	Kanai Y.;	
RT	"Cloning and characterization of a human system L amino acid	
RT	transporter.";	
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Placenta;	
RX	MEDLINE=99160855; PubMed=10049700;	
RA	Prasad P.D., Wang H., Huang W., Kekuda R., Rajan D.P., Leibach F.H.,	
RA	Ganapathy V.;	
RT	"Human LAT1, a subunit of system L amino acid transporter: molecular	
RT	cloning and transport function.";	
RL	Biochem. Biophys. Res. Commun. 255:283-288(1999).	
RN	[4]	
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.	
RX	MEDLINE=99172172; PubMed=10072483;	
RA	Tsurudome M., Ito M., Takebayashi S., Okumura K., Nishio M.,	
RA	Kawano M., Kusagawa S., Komada H., Ito Y.;	
RT	"Primary structure of the light chain of fusion regulatory protein-	
RT	1/CD98/4f2 predicts a protein with multiple transmembrane domains that	
RL	is almost identical to the amino acid transporter El6.";	
RL	J. Immunol. 162:2462-2466(1999).	
RN	[5]	
RP	SEQUENCE FROM N.A.	
RA	Minato N., Iwai K., Takizawa C., Nakamura E.;	
RT	"Human 4f2 light chain: amino acid transporter.";	
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.	
RN	[6]	
RP	SEQUENCE OF 267-507 FROM N.A.	
RC	TISSUE=peripheral blood lymphocytes;	
RX	MEDLINE=92283834; PubMed=1597461;	
RA	Gaugitsch H.W., Prieschl E.E., Kalthoff F., Huber N.E.,	
RA	Baumruker T.;	
RT	"A novel transiently expressed, integral membrane protein linked to	
RT	cell activation. Molecular cloning via the rapid degradation signal	
RT	AJUA.";	
RL	J. Biol. Chem. 267:11267-11273(1992).	
CC	-1- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE	
CC	NEUTRAL AMINO ACIDS. INVOLVED IN CELLULAR AMINO ACID UPTAKE.	
CC	-1- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID	
CC	TRANSPORTER PROTEIN SLC3A2/4F2HC.	
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).	
CC	-1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN ADULT LUNG AND LIVER,	
CC	AND IS ALSO EXPRESSED IN BRAIN, THYMUS, RETINA AND SOME OTHER	
CC	TISSUES.	
CC	-1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID	
CC	PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)	
CC	SUBFAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/)	

QY 1681 ACCAGACATACAAATAATCTGGAGTGTACAGAGAA 1722
 ||| ||| :
 Db 493 ThrValLeuCysGlnLysLeuMetGlnValProGlnGlu 506

RESULT 4
 LAT1L_RAT STANDARD; PRT; 512 AA.

AC Q63016; Q9QWL4;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Large neutral amino acids transporter small subunit 1 (L-type amino
 acid transporter 1) (4F2 light chain) (4F2 LC) (4F2LC) (Integral
 membrane protein E16) (TAL protein).
 DE SLC7A5 OR MPE16 OR TAL.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 OX [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=98395066; PubMed=9726963;
 RA Kanai Y., Segawa H., Miyamoto K., Uchino H., Takeda E., Endou H.;
 RA "Expression cloning and characterization of a transporter for large
 RT neutral amino acids activated by the heavy chain of 4F2 antigen
 RT (CD98).";
 RT J. Biol. Chem. 273:23629-23632(1998).
 RL [2]
 RP SEQUENCE OF 272-512 FROM N.A.
 RP TISSUE=Hepatoma;
 RX MEDLINE=95171385; PubMed=7532544;
 RA Sang J., Lim Y.P., Panzica M., Thompson N.L.;
 RA "TAL, a highly conserved oncofetal complementary DNA from rat
 RT hepatoma, encodes an integral membrane protein associated with liver
 RT development, carcinogenesis, and cell activation.";
 RT Cancer Res. 55:1152-1159(1995).
 RL -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE
 CC NEUTRAL AMINO ACIDS. INVOLVED IN CELLULAR AMINO ACID UPTAKE.
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
 CC TRANSPORT PROTEIN SLC3A2/4F2HC.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- TISSUE SPECIFICITY: EXPRESSED HEPATOMA BUT NOT IN NORMAL LIVER.
 CC ALSO EXPRESSED IN PLACENTA, TESTIS, BRAIN, OVARY, SPLEEN, MAMMARY
 CC GLAND, AND UTERUS.
 CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
 CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
 CC SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; AA015432; BAA33035.1; -
 DR EMBL; U00995; AAA74411.1; -
 DR InterPro: IPR002293; AA/rel_prmease1.
 DR InterPro: IPR004760; LAA_transport.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 DR TIGRFAMs: TIGR00911; ZAO308; 1.
 KW Transport; Amino-acid transport; Transmembrane.
 FT TRANSMEM 51 71 POTENTIAL.
 FT TRANSMEM 85 105 POTENTIAL.
 FT TRANSMEM 121 141 POTENTIAL.
 FT TRANSMEM 147 167 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT TRANSMEM 200 220 POTENTIAL.
 FT TRANSMEM 248 268 POTENTIAL.
 FT TRANSMEM 279 299 POTENTIAL.

Thu Apr 17 09:48:26 2003

us-09-667-170a-440.n2p.rsp

111 ILeSerlySerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAla 130
598 TTTGTAGAGCTCGGGTGGAACTCCATAAATACGCCCTCGAGCTACTCTGTGATATCC 657
131 PheLeuLysLeuTrpIleGluLeuLeuIleArgProSerSerGlnTyrIleValAla 150
658 CTGCATTGGAGCGCTACATCTCGAACCACTTTTATTCATGTAATCCCTGAACAT 717
151 LeuValPheAlaThrTyrLeuLeuLysProValPheProThrCysProValProGluGlu 170
718 GCGATCAAGCTCATACAGCTGTGGGCTAACTAGTGTAGTGTCTCTAAATGATGAT 777
171 AlaAlaLysLeuValAlaLysLeuLysValLeuLeuLeuThrAlaValAsnCysTyrSer 190
778 GTACGCTGGAGCGCGGATCCAGATTTCTTAACCTTTTGAAGCTCAGACAAATCTG 837
191 ValLysAlaAlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAla 210
838 ATAATTATAGTCCCTGGAGTATGACCTAATTAATTAATTAATTAATTAATTAAT 885
211 LeuIleLeuLeuLeuGlyPheIleGlnMetGlyLysAspMetGlyGlnGlyAspAlaSer 230
886 AACTTT-----AAAGCGCTCTTTCAGGAGAGATTAAGTATTCAGCGGTTGCCACTG 939
231 AsnLeuGlnGlnLysLeuSerPheGluGlyThrAsnLeuAspValGlyAsnIleValLeu 250
940 GCTTTTATTATGAATGATGATATGCTGCTGGCTGCTTTTACCTCACTTTGTTACTGAA 999
251 AlaLeuTyrSerGlyLeuPheAlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGlu 270
1000 GAAGTAGAAACCCCTGAAACCACTCCCTTCGATATGATATGATATGATATGATATG 1059
271 GluMetIleAsnProTyrArgAsnLeuProLeuAlaIleIleSerLeuProIleVal 290
1060 ACCATTGGCTATGCTGACAAATGTCCTACTTACGACCATTAATGCTGAGAGCTG 1119
291 ThrLeuValTyrValLeuThrAsnLeuAlaTyrPheThrThrLeuSerThrAsnGlnMet 310
1120 CTGCTTTCATATGCTGAGTGCAGTACCTTTCTGAGCGCTACTGGGAAATTTCTCATTA 1179
311 LeuThrSerGluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSerTrp 330
1180 GCAGTTCCGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239
331 IleIleProValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThr 350
1240 GTCTCCAGGTTATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1299
351 SerSerArgLeuPhePheValGlySerArgGlyHisLeuProSerValLeuSerMet 370
1300 ATTCATGCTCCGCAAGCACACTCTCTACCACTGTTATTTGTTGCTGCTGCTGCTGCT 1359
371 IleHisProGlnLeuLeuThrProValProSerLeuValPheThrCysIleMetThrLeu 390
1360 ATAATGCTCTCTCTGAGACCTCGACACTCTTTTGAATTTCTCAGTTTCCGAGTGG 1419
391 MetTyrAlaPheSerArgAspIlePheSerIleIleAsnPhePheSerPheAsnTrp 410
1420 CTTTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1479
411 LeuCysValAlaAlaIleIleGlyMetMetTrpLeuArgPheLysLysProGluLeu 430
1480 CATGCTGCTTTCAGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1539
431 GluArgProIleLysValAsnLeuAlaLeuProValPhePheIleLeuAlaCysLeuPhe 450
1540 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1599
451 LeuIleAlaValSerPheThrLysThrPrometGluCysGlyIleGlyPheAlaIleIle 470
1600 CTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1659

Db 471 LeuSerGlyLeuProValTyrPhePheGlyValTyrTrpLysAsnLysProLysTrpIle 490
QY 1660 AGAATAATGTTCAGAGAAATAACACAGAACATTAACAATAATACATGGAAGTTGTACCA 1719
Db 491 LeuGlnAlaIlePheSerValThrValLeuCysGlnLysLeuMetGlnValValProGln 510
QY 1720 GAA 1722
Db 511 Glu 511
RESULT 6
YLAL_HUMAN STANDARD; PRT: 511 AA.
AC Q9UM01; Q9P2V5; O95984;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Y+L amino acid transporter 1 (Y(+)-L-type amino acid transporter 1)
DE (Y+LAT-1) (Y+LAT1) (Monocyte amino acid permease 2) (MOP-2).
GN SLC7A7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT LPI ARG-334.
RX MEDLINE=99047611; PubMed=9829974;
RA Torrents D., Estevez R., Pineda M., Fernandez E., Lloberas J.,
RA Shi Y.-B., Zorzano A., Palacin M.;
RT "Identification and characterization of a membrane protein (Y+L amino
RT acid transporter-1) that associates with 4F2hc to encode the amino
RT acid transport activity Y+L. A candidate gene for lysinuric protein
RT intolerance.";
RL J. Biol. Chem. 273:32437-32445(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=99094891; PubMed=9878049;
RA Pfeiffer R., Kossler G., Spindler B., Meier C., Kuhn L., Verrey F.;
RT "Amino acid transport of Y+L-type by heterodimers of 4F2hc/CD98 and
RT members of the glycoprotein-associated amino acid transporter
RT family.";
RL EMBO J. 18:49-57(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=99178268; PubMed=10080183;
RA Borsani G., Bassi M.T., Sperandio M.P., De Grandi A., Buoninconti A.,
RA Riboni M., Manzoni M., Incerti B., Pepe A., Andria G., Ballabio A.,
RA Sebastio G.;
RT "SLC7A7, encoding a putative permease-related protein, is mutated in
RT patients with lysinuric protein intolerance.";
RL Nat. Genet. 21:297-301(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Fukasawa Y., Segawa H., Endou H., Kanai Y.;
RT "Characterization of a human system Y+L amino acid transporters.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20202143; PubMed=10737982;
RA Noguchi A., Shoji Y., Koizumi A., Takahashi T., Shoji Y.,
RA Matsumori M., Kayo T., Ohata T., Wada Y., Yoshimura I., Maisawa S.,
RA Konishi M., Takasago Y., Takada G.;
RT "SLC7A7 genomic structure and novel variants in three Japanese
RT lysinuric protein intolerance families.";
RL Hum. Mutat. 15:367-372(2000).
RN [6]
RP SEQUENCE FROM N.A.
RA Takayama K., Yoshimoto M.;
RT "Molecular and biological characterization of a novel monocyte amino
RT acid permease, MOP-2.";

	850	CCTGGAGTTATGACAGTAATTAAAGTCAACGCGAAGACTTTAAAGCAGCGCTTTTCAGGA	909
QY		: : : : : : : : : : :	
Df	201	AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGluGly	220
QY	910	AGAGATTCAAAGTATTACGGGGTTGCCCATGTCCGCTTTTTATTATGGAATGTATGCATATGCT	969
Df	221	SerSerPheAlaValGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer	240
QY	970	GGCTGGTTTACCTCAACTTTGTCTACAGAAGTAGAAAACCCCTGAAAAAACCATCCC	1029
Df	241	GlyTrpAspThrLeuAsnTyrValThrGluGluIleLysAsnProGluArgAsnLeuPro	260
QY	1030	CTTGCAATATGATATCATCGAGCCATTGTACACATTGGCTATGTCTCTCACAAATGTGGCC	1089
Df	261	LeuSerIleGlyIleSerMetProIleValThrIleIleTyrIleLeuThrAsnValala	280
QY	1090	TACTTTAGGACCATTAAATGCTGGAGAGCTGCTGCTTTCAATGCGAGTGGCAGTGACCTTT	1149
Df	281	TyrtYrThrValLeuAspMetArgAspIleLeuAlaSerAspAlaValAlaValThrPhe	300
QY	1150	TCGTAGCGGCTACTGGGAAATTTCTCATAGCACTGCCGATCTTTGTTCGCTCTCPGCG	1209
Df	301	AlaAspGlnIlePheGlyIlePheAsnTrpIleIleProLeuSerValAlaLeuSerCys	320
QY	1210	TTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTATTCTATCTATGCTGCGTCTCGA	1269
Df	321	PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuPheValGlySerArg	340
QY	1270	GAGGTTCACCTCCAGAATACTCTCCATGATTCATCGCGAAGCACACTCCTCTACCA	1329
Df	341	GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro	360
QY	1330	GCTGTATTATGTTTGCACCCCTTTGCAATGATATGCTCTCTCGAGACCTCGACAGT	1389
Df	361	SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln	380
QY	1390	CTTTTGAATTTCTCTCAGTTTTCGCCAGGTGGCTTTTATTGGCTGGCAGTGTCTGGGCTG	1449
Df	381	LeuIleAsnTyrTyrSerPheSerTyrTrpPhePheValGlyLeuSerIleValGlyGln	400
QY	1450	ATTTATCTTCGATACAAATGCCAGATATGATCGTCTTTCAGSGTCCACCTGTTTCATC	1509
Df	401	LeuTyrLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerValPhePhe	420
QY	1510	CCACCTTTGTTTTCCTTCACATGCCTCTTCATGGTGGCTTTCCCTCTATTCCGACCCA	1569
Df	421	ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr	440
QY	1570	TTTAGTACAGGGATGGCTTCGTCATCACTCTGACTGGAGTCCCTGGGTATTATCTCTTT	1629
Df	441	IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIle	460
QY	1630	ATT-----ATATGGGACAAGAACCCAGGTGGTTGATAATATATCTCAGAGAAATTAACC	1683
Df	461	IleArgValProGluHisLysArgProLeuTyrLeuArgArgIleValcglySerAlaThr	480
QY	1684	AGAATTTACAAATAATA---CTGGAGTGTGTACACAGAGAAGAT	1725
Df	481	ArgTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetasp	495
<hr/>			
	RESULT 7		
	LAT2_RAT	STANDARD;	PRT; 533 AA.
ID	LAT2_RAT		
AC	Q9WVR6;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Large neutral amino acids transporter small subunit 2 (L-type amino acid transporter 2).		
GN	SLC7A8 OR LAT2.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		

NCBI_TaxID=101116;

[1]

SEQUENCE FROM N.A.

TISSUE=Small intestine;

MEDLINE=99321902; PubMed=10391916;

Segawa H., Fukasawa Y., Miyamoto K., Takeda E., Endou H., Kanai Y.;

"Identification and functional characterization of a Na+-independent

neutral amino acid transporter with broad substrate selectivity.";

J. Biol. Chem. 274:19745-19751(1999).

-!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE

NEUTRAL AMINO ACIDS. HAS HIGHER AFFINITY FOR L-PHENYLALANINE THAN

LATI. L-ALANINE IS TRANSPORTED AT PHYSIOLOGICAL CONCENTRATIONS.

PLAYS A ROLE IN BASOLATERAL (RE)ABSORPTION OF NEUTRAL AMINO ACIDS.

-!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID

TRANSPORT PROTEIN SLC3A2/4F2HC.

-!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

-!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID

PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)

SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

or send an email to license@isb-sib.ch).

EMBL; AB024400; BAA82517.1; -

InterPro; IPR002293; AA/rei_pmeasel.

InterPro; IPR004760; L_AA.transport.

InterPro; IPR004841; Permease.

Pfam; PF00324; aa_permeases; 1.

TIGRfams; TIGR00911; 2A0308; 1.

DR DR TRANSPORT; Amino acid transport; Transmembrane.

KW TRANSMEM 41 61 POTENTIAL.

FT TRANSMEM 73 93 POTENTIAL.

FT TRANSMEM 114 134 POTENTIAL.

FT TRANSMEM 156 176 POTENTIAL.

FT TRANSMEM 190 210 POTENTIAL.

FT TRANSMEM 232 252 POTENTIAL.

FT TRANSMEM 269 289 POTENTIAL.

FT TRANSMEM 311 331 POTENTIAL.

FT TRANSMEM 363 383 POTENTIAL.

FT TRANSMEM 389 409 POTENTIAL.

FT TRANSMEM 423 443 POTENTIAL.

FT TRANSMEM 448 468 POTENTIAL.

SQ SEQUENCE 533 AA; 58190 MW; 99479DB6DA59DF0 CRC64;

Alignment Scores:

Pred. No.: 2,09e-79 Length: 533 -

Score: 1108.00 Matches: 224

Percent Similarity: 63.04% Conservative: 100

Best Local Similarity: 43.58% Mismatches: 168

Query Match: 26.89% Indels: 22

DB: 1 Gaps: 5

US-09-667-170A-440 (1-2239) x LAT2_RAT (1-533)

OY 202 ATCTCCAAAGAGGTTACCTGCAGGAAATGTTAAGCGGAGCTGCCTTCCTGGCAAC 261

Db 1 MetGluLysGlyThrArgGlnArgAsnThrAlaLysAsnHisProAspArgGlySer 20

OY 262 AAGGAGCCACTGGCGAGGAGCGCTTTTCAGGAGAGACGCCCTTTTCAGGAAGAGAGAGCC 321

Db 21 AspThrSerProGluAlaGluAlaSerSerGlyGlyGlyCly----- 34

OY 322 TTTTCAGGAAGAGAGAGAAGTGCAGCTCAGAGAGAAAGTCACTTTACTGAGGGAGCTCC 381

Db 35 -----ValAlaLeuLysLysGluIleGlyLeuValserAlaCysGly 48

OY 382 ATTATCAATGGCACCATCAATTGGAGCAGGAATCTCATCTCTCCTAAGCGCGTGCCTCAG 441

Db 35 -----ValAlaLeuLysLysGluIleGlyLeuValserAlaCysGly 48

OY 382 ATTATCAATGGCACCATCAATTGGAGCAGGAATCTCATCTCTCCTAAGCGCGTGCCTCAG 441

Db 35 -----ValAlaLeuLysLysGluIleGlyLeuValserAlaCysGly 48

Db 49 IieIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 68
 QY 442 AACACGGCAGCGTGGGATGCTCTGTACCATCTGAGCGGTGTGGGGTCCCTGCTCACTA 501
 Db 69 AsnAlaGlySerValGlyLeuAlaLeuAlaLeuValTrpIleValThrGlyValIleThrAla 88
 QY 502 TTGGAGCTTTGCTTATGCTGAATTTGGGAACAACCTATAAGAAATCTGAGGTCAATAC 561
 Db 89 ValGlyAlaLeuCysTyrAlaGluLeuGlyValTrpIleProLysSerGlyIleAspTyr 108
 QY 562 ACATATATTTTGGAGCTTTGGTCCCAATACCACTTTTGTACAGTCTGGGTGGAACTC 621
 Db 109 SerTyrValLysAspIlePheGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 128
 QY 622 CTCATAATACCCCTGCGAGCTACTGCTGATATCCCTGCGATTTGGAGCTCAATTCG 681
 Db 129 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 148
 QY 682 GAACCATTTTATTCATGTAATCCCTGAATCCCTGAATTCGATCAAGCTCAATACAGCTG 741
 Db 149 GlnProLeuPheProThrCysPheProGluSerGlyLeuArgLeuLeuAlaIle 168
 QY 742 GGCATAACTGTAGTATGCTTAAATAGCATGCTGAGTGGAGCCCGGATCCAG 801
 Db 169 CysLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 188
 QY 802 ATTTCTTAACTTTTGCAGCTCACAGCAATTCATATATATAGTCCCTGGAGTTATG 861
 Db 189 AspIlePheThrAlaGlyLysLeuLeuAlaLeuIleIleMetGlyValVal 208
 QY 862 CAGCTAATTAAGCTCAACGCGAAGCTTT-----AAAGAGCGCTTTTCA 906
 Db 209 GlnIleCysLysGlyGlu-----PhePheTrpLeuGluProLysAsnAlaPheGlu 225
 QY 907 GGA---AGAGATCAAGTATTACGGGTGGTCCACTGCTTTTATATGGAATGATGCA 963
 Db 226 AsnPheGlnGluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAla 245
 QY 964 TATGCTGGCTGTTTACCTCACTTGTACTGAAGAGTAGAAGAACCCCTGAAAAACC 1023
 Db 246 TyrGlyGlyTrpAsnPheLeuAsnTyrValThrGluGluLeuValAlaSerProTyrLysAsn 265
 QY 1024 ATCCCTTTCGAATATGATATTCATGCGCATTCACCATTCGCTATGCTGACAAAT 1083
 Db 266 LeuProArgAlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsn 285
 QY 1084 GTGGCTACTTTACGACCATTAATGCTGAGGAGTGTCTTCAATGAGTGGCAGTG 1143
 Db 286 IleAlaTyrValThrAlaMetSerProGlnGluLeuLeuAlaSerAsnAlaValAlaVal 305
 QY 1144 ACCTTTTCTGAGCGCTACTCGGAATTCCTCATTAGCATTCGATCTTTGTGGCCTC 1203
 Db 306 ThrPheGlyGlyLysLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeu 325
 QY 1204 TCCTGCTTGGCTCCATGAAGCGGTGTGTCTGCTCCAGGTATATCTATGTTGGG 1263
 Db 326 SerThrPheGlyGlyValAsnGlySerLeuPheThrSerSerArgLeuPheAlaGly 345
 QY 1264 TCTCAGAGGCTCACCCTCCAGAAATCCTCCATGATTCATCCGCAAGCAGCATCTCT 1323
 Db 346 AlaArgGluGlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrPro 365
 QY 1324 CTACAGCTGTATTGTTTGGACCCCTTTGCACATCATATGCTCTCTCTGGAGACCTC 1383
 Db 366 IleProAlaLeuLeuPheThrCysLeuSerThrLeuLeuMetLeuValThrSerAspMet 385
 QY 1384 GACAGCTTTTGAATTTCCAGTTTGGCAGGTGCTTTTATGGGTGGCAGTGTCT 1443
 Db 386 TyrThrLeuIleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValValAla 405
 QY 1444 GGGCTGATTTATCTCGATACAATCCCGATATGTCATGCTCTTCAAGGTGCCACTC 1503
 Db 406 GlyGlnIleValLeuArgTrpLysProAspIleProArgProIleLysIleSerLeu 425

QY 1504 TTCATCCAGCTTTGTTTTCCTTCATGCTTCATGTTGCCCTTCCCTCTATTTCG 1563
 Db 426 LeuPheProIleIleTyrLeuLeuPheTrpAlaPheLeuLeuIlePheSerLeuTrpSer 445
 QY 1564 GACCCATTAGTACAGGATTCGCTTCATCATCTGCTGAGTCCCTGGGTATAT 1623
 Db 446 GluProValValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPhe 465
 QY 1624 CTCCTTTATATATGGGACAAAGAACCCAGGTGTTTACAATAATGTCAGAGAAATAACC 1683
 Db 466 LeuGlyValTyrTrpGlnHisLysProLysCysPheAsnAspPheIleGluSerLeuThr 485
 QY 1684 AGAATTTACAAATAATCTGGAAGTTGTA---CCAGAGAA 1722
 Db 486 LeuValSerGlnLysMetCysValValValTyrProGlnGlu 499

RESULT 8
 LAT2_MOUSE
 ID LAT2_MOUSE STANDARD; PRT; 531 AA.
 AC Q9XW9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Large neutral amino acids transporter small subunit 2 (L-type amino acid transporter 2).
 DE SLC7A8 OR LAT2.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=20044753; PubMed=10574970;
 RA Rossier G., Meier C., Bauch C., Summa V., Sordat B., Verrey F., Kuhn L.C.;
 RT "LAT2, a new basolateral 4F2hc/CD98-associated amino acid transporter of kidney and intestine.";
 RL J. Biol. Chem. 274:34948-34954(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20079165; PubMed=10610726;
 RA Bassi M.T., Sperandio M.P., Incerti B., Bulfone A., Pepe A., Surace E.M., Gattuso C., de Grandi A., Buoinconti A., Riboni M., Manzoni M., Andria G., Ballabio A., Borsani G., Sebastio G.;
 RT "SLC7A8, a gene mapping within the lysinuric protein intolerance critical region, encodes a new member of the glycoprotein-associated amino acid transporter family.";
 RL Genomics 62:297-303(1999).
 CC -!- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE NEUTRAL AMINO ACIDS. HAS HIGHER AFFINITY FOR L-PHENYLALANINE THAN LAT1. L-ALANINE IS TRANSPORTED AT PHYSIOLOGICAL CONCENTRATIONS. PLAYS A ROLE IN BASOLATERAL (RE)ABSORPTION OF NEUTRAL AMINO ACIDS.
 CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID TRANSPORT PROTEIN SLC3A2/4F2HC.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- TISSUE SPECIFICITY: MAINLY EXPRESSED IN KIDNEY AND INTESTINE.
 CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT) SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/annouce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF171668; AAF20380.1; -
 CC EMBL; Y19022; CAB69072.1; -
 CC MGI; 1355923; Slc7a8.
 CC -----
 CC DR
 CC DR

PS-09-6667-170A-440 (1-2239) X LAT2_MOUSE (1-531)

[illegible]

RESULT 9

LAT2_HUMAN

LAT2_HUMAN	STANDARD;	PRT;	535 AA.
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			

LIAI2_HUMAN
Q9UHT5; Q9Y445; Q9UKQ6; Q9UKQ7; Q9UKQ8;

Q90H15; Q91443; Q90K87; Q90K86;
16-OCT-2001 (Rel. 40, Created)

16-OCT-2001 (REL. 40, Creaca)
16-OCT-2001 (REL. 40, Last sequence update)

DT	16-OCT-2001 (Rel. 40,	17-JUN-2002 (Rel. 41,
	Last sequence update)	Last annotation update)

15-JUN-2002 (REL. 41, Last annotation update)

Large neutral amino acids transporter small

DE acid transporter 2) (hLAT2).

CN SLC7A8 OR LAT2.

Homo sapiens (Human).

05. Homo sapiens (human);
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata

Mammalia: Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homininae; Homo

NCBI TaxID=9606;

NCBI_TaxID=9606;

RN [1] REFERENCES FROM N A

RP SEQUENCE FROM N.A.

RX MEDLINE=20044753; PubMed=10574970;
 RA Rossier G., Meier C., Bauch C., Summa V., Sordat B., Verrey F.,
 RA Kuhn L.C.;
 RT "LAT2, a new basolateral 4F2hc/CD98-associated amino acid transporter
 of kidney and intestine.";
 RL J. Biol. Chem. 274:34948-34954(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=99178268; PubMed=10080183;
 RA Borsani G., Bassi M.T., Sperandio M.P., De Grandi A., Buoninconti A.,
 RA Riboni M., Manzoni M., Incerti B., Pepe A., Andria G., Ballabio A.,
 RA Sebastio G.;
 RT "SLC7A7, encoding a putative permease-related protein, is mutated in
 patients with lysinuric protein intolerance.";
 RL Nat. Genet. 21:297-301(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99321901; PubMed=10391915;
 RA Pineda M., Fernandez E., Torrents D., Estevez R., Lopez C., Camps M.,
 RA Lloberas J., Zorzano A., Palacin M.;
 RT "Identification of a membrane protein, LAT-2, that co-expresses with
 4F2 heavy chain, an L-type amino acid transport activity with broad
 specificity for small and large zwitterionic amino acids.";
 RL J. Biol. Chem. 274:19738-19744(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Kim D.-K., Fukushima J., Segawa H., Cha S.-H., Chairoungdua A.,
 RA Matsuo H., Kim J.-Y., Goya T., Endou H., Kanai Y.;
 RT "Molecular cloning and characterization of a human Na⁺-independent
 neutral amino acid transporter expressed in the basolateral membrane
 of renal proximal tubule.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SODIUM-INDEPENDENT, HIGH-AFFINITY TRANSPORT OF LARGE
 NEUTRAL AMINO ACIDS. HAS HIGHER AFFINITY FOR L-PHENYLALANINE THAN
 LAT1. L-ALANINE IS TRANSPORTED AT PHYSIOLOGICAL CONCENTRATIONS.
 CC PLAYS A ROLE IN BASOLATERAL (RE)ABSORPTION OF NEUTRAL AMINO ACIDS.
 CC SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID
 TRANSPORT PROTEIN SLC3A2/4F2HC.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- TISSUE SPECIFICITY: MAINLY EXPRESSED IN KIDNEY AND INTESTINE.
 CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID
 PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAAT)
 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF171669; AAF20381.1; -
 CC EMBL; Y18483; CAB40137.1; -
 CC EMBL; AF135830; AAF05697.1; -
 CC EMBL; AF135829; AAF05696.1; -
 CC EMBL; AF135828; AAF05695.1; -
 CC EMBL; AB037669; BAB21519.1; -
 CC EMBL; HGNC:11066; SLC7A8.
 CC MIM; 604235; -
 CC InterPro; IPR002293; AA/reli_pmease1.
 CC InterPro; IPR004760; LAA_transport.
 CC InterPro; IPR004841; Permease.
 CC Pfam; PF00324; aa_permeases; 1.
 CC TIGRFAMs; TIGR00911; 2A0308; 1.
 CC Transport; Amino-acid transport; Transmembrane.
 KW TRANSMEM 40 60 POTENTIAL.
 FT TRANSMEM 72 92 POTENTIAL.
 FT TRANSMEM 113 133 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.

FT TRANSMEM 231 251 POTENTIAL.
 FT TRANSMEM 268 288 POTENTIAL.
 FT TRANSMEM 310 330 POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 388 408 POTENTIAL.
 FT TRANSMEM 424 444 POTENTIAL.
 FT TRANSMEM 447 467 POTENTIAL.
 FT CONFLICT 225 225 N -> D (IN REF. 2).
 FT CONFLICT 401 401 V -> G (IN REF. 2).
 FT CONFLICT 503 503 G -> R (IN REF. 3).
 SQ SEQUENCE 535 AA; 58381 MW; AC129146353F1E47 CRC64;
 Alignment Scores:
 Pred. No.: 6.8e-79 Length: 535
 Score: 1101.50 Matches: 220
 Percent Similarity: 65.52% Conservative: 103
 Best Local Similarity: 44.62% Mismatches: 161
 Query Match: 26.74% Indels: 9
 DB: 1 Gaps: 4
 US-09-667-170a-440 (1-2239) x LAT2_HUMAN (1-535)
 QY 262 AAGGAGCCACCTGGCAGGAGCGCTTTTCAGGAGAGAGCCCTTTTCAGGAGAGAGAGCGC 321
 DB 13 LysLysHisProGly-----GlyGlyGlySerAspAlaSerProGluAla 27
 QY 322 TTTTCAGGAGAGAGAAAGTGCAGCTGAGAGAGAAAGTCACTTTACTGAGGGAGTCTCC 381
 DB 28 GlySerGlyGlyGlyValAlaLeuLysLysGluLeuValSerAlaCysGly 47
 QY 382 ATTATCATTCGGCACCATTGAGGAGGAGAAATCTTCATCTCTCAAGGGCGTCTCCAG 441
 DB 48 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 67
 QY 442 AACACGGGACGCTGGCATGCTCTGACCATCTGCACGCTGGCGGTCTCTGCATCA 501
 DB 68 AsnAlaGlySerValGlyLeuAlaLeuIleValIleValThrGlyPheIleThrVal 87
 QY 502 TTTGGAGCTTTGCTTATGCTCAATTTGGAAACAACATATAAGAAATCTGGAGTCTTAC 561
 DB 88 ValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr 107
 QY 562 ACATATATTTTGAAGCTTTTGGTCCATCCATACCAGCTTTTGACGACTCTGGGTGGAAC 621
 DB 108 SerTyrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 127
 QY 622 CTCATATAACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGAGCGCTACATTCG 681
 DB 128 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 147
 QY 682 GAACCATTTTATTCATGCAATCCCTGAATCGGATCGGATCGGATCAAGCTGATG 741
 DB 148 GlnProLeuPheProThrCysPheProGluSerGlyLeuArgLeuAlaAlaIle 167
 QY 742 GCATATACTGTAGTGTGCTCTTAATAGCATGCTGACGTGACGTGAGCGCCCGCATCCAG 801
 DB 168 CysLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 187
 QY 802 ATTTCTTAACCTTTTGAAGCTCACAGCAATTCGATTAATATAGTCCCTGGAGTTATG 861
 DB 188 AspIlePheThrAlaGlyLysLeuAlaLeuIleIleMetGlyIleVal 207
 QY 862 CAGCTAAATTAAGGTCAA-----ACGAGAACCTTTAAAGACGCTTTTCAGGA---AGA 912
 DB 208 GlnIleCysLysGlyGlyTyrPheTrpLeuGluProLysAsnAlaPheGluAsnPheGln 227
 QY 913 GATTCACGATTACCGGCTTCCCTGCTTTTATTATGAAATGATGATCATATGCTGGC 972
 DB 228 GluProAspIleGlyLeuValAlaAlaPheLeuGlnGlySerPheAlaTyrGlyGly 247
 QY 973 TGGTTTTACCTCAACTTTGTTACTGAAGAGTAGAAACCCCTGAAAAACCATTCGCCCTT 1032
 DB 248 TrpAsnPheLeuAsnTyrValThrGluLeuValAspProTyrLysAsnLeuProArg 267

FT TRANSMEM 375 395 POTENTIAL.
 FT TRANSMEM 410 430 POTENTIAL.
 FT TRANSMEM 435 455 POTENTIAL.
 FT VARIANT 105 105 G -> R (IN CSNU3).
 FT VARIANT 170 170 /FTID-VAR_010256.
 FT VARIANT 182 182 V -> M (IN CSNU3).
 FT VARIANT 182 182 /FTID-VAR_010257.
 FT VARIANT 195 195 A -> T (IN CSNU3).
 FT VARIANT 195 195 /FTID-VAR_010258.
 FT VARIANT 259 259 G -> R (IN CSNU3).
 FT VARIANT 259 259 /FTID-VAR_010259.
 FT MUTAGEN 170 170 G -> R (IN CSNU3).
 FT MUTAGEN 170 170 /FTID-VAR_010260.
 FT CONFLICT 52 52 V->M: LOSS OF AMINO ACID TRANSPORT
 FT CONFLICT 52 52 ACTIVITY.
 SQ SEQUENCE 487 AA; 53481 MW; EF2C30DDE15594F1 CRC64;

Alignment Scores:
 Pred. No.: 7,43e-77 Length: 487
 Score: 1075.50 Matches: 207
 Percent Similarity: 66.16% Conservative: 98
 Best Local Similarity: 44.90% Mismatches: 153
 Query Match: 26.10% Indels: 3
 DB: 1 Gaps: 2

US-09-667-170A-440 (1-2239) x BAT1_HUMAN (1-487)

QY 346 CTGAAGAGGAAGTCACTTTACTAGGAGGAGTCTCCATTATCATATGCGACCATCATGGA 405
 Db 26 LeuGlnLysLeuLeuGlyLeuLeuSerGlyLeuSerIleValGlyThrIleGly 45
 QY 406 GCAGGAATCTTCACTCTCTAAGGGGTGCTCCAGAACACGGGCGGTGGGCATGTCT 465
 Db 46 SerGlyIlePheValSerProLysSerValLeuSerAsnThrGluAlaValGlyProCys 65
 QY 466 CTGACCATCTGGACGGTGTGGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585
 Db 66 LeuIleThrAlaAlaCysGlyValLeuAlaThrLeuGlyAlaLeuCysPheAlaGlu 85
 QY 526 TTGGGAACAACATATAAGAAATCTGGAGGTCAATACATATATTTTGAAGTCTTTGGT 585
 Db 86 LeuGlyThrMetIleThrLysSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 105
 QY 586 CCATTACCAAGCTTTGTAGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 645
 Db 106 ProIleProAlaThrLeuPheSerTrpAlaSerLeuIleValIleLysProThrSerPhe 125
 QY 646 GCTGTGATATCCCTGGCATTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
 Db 126 AlaIleIleCysLeuSerPheSerGlyTrpValCysAlaProPheThrValGlyCysLys 145
 QY 706 ATCCCTGAACCTTGGCATCAAGCTCATACAGCTGTGGGCAATACTGTAGTGATGCTCTA 765
 Db 146 ProProGlnIleValValCysLeuAlaAlaAlaAlaIleLeuPheIleSerThrVal 165
 QY 766 AATGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 825
 Db 166 AsnSerLeuSerValArgLeuGlySerTrpValGlnAsnIlePheThrAlaAlaLysLeu 185
 QY 826 ACAGCAATCTGTGAATTAATAGTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 885
 Db 186 ValIleValAlaIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 205
 QY 886 AACTTTAAGACGCTTTTTCAGGAAGAGATTCAGATTAATACGGGTGGTGGTGGTGGTGGT 945
 Db 206 AsnPheAsnSerPheGluGlyAlaGlnLeuSerValIleAlaIleSerLeuAlaPhe 225
 QY 946 TATTATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1005
 Db 226 TyrAsnGlyLeuThrAlaThrAspGlyTrpAsnGlnLeuAsnThrIleThrGluGluLeu 245
 QY 1006 GAAACCCCTGAAAAACCATTCCTCCCTTGCATATGATATGATATCATCATGCCATTGT 1065

Db 246 ArgAsnProTyrArgAsnLeuProLeuAlaIleIleIleIleIleIleIleIleIleIleIle 265
 QY 1066 GGCTATGCTGTCGCAAAATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCT 1125
 Db 266 CystyrIleLeuMetAsnValSerTyrPheThrValMetThrAlaThrGluLeuLeuGln 285
 QY 1126 TCAAATGAGTGGCAGTGGCTTTCTGACGGGCTACTGGGAAATTTCTCATTTAGCAGTT 1185
 Db 286 SerGlnAlaValAlaValThrPheGlyAspArgValLeuTyrProAlaSerTrpIleVal 305
 QY 1186 CCGATCTTTGTGCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1245
 Db 306 ProLeuPheValAlaPheSerThrIleGlyAlaAsnGlyThrCysPheThrAlaGly 325
 QY 1246 AGGTATTCTATGCTGGTCTCGAGAGGTACCTTCCAGAAATCCCTCCCATGATTCAT 1305
 Db 326 ArgLeuIleThrValAlaGlyArgGlyHisMetLeuLysValLeuSerTrpIleSer 345
 QY 1306 GTCCGCAAGCACACTCTCTTACCAGCTGTTATTGTTTTCACCCCTTGGCAATGATAATG 1365
 Db 346 ValArgArgLeuThrProAlaProAlaIleIlePheThrGlyIleIleAlaThrIleTyr 365
 QY 1366 CTCTTCTCTGGAGACCTCGACAGTCTTTTCAATTTCTCAGTTTTCAGGAGTGGCTTTT 1425
 Db 366 IleIleProGlyAspIleAsnSerLeuValAsnTyrPheSerPheAlaIleThrLeuPhe 385
 QY 1426 ATTGGCTGGCAGTCTGCTGGCTGATTTATCTTCATACAAATGCCAGATATGATCGT 1485
 Db 386 TyrGlyLeuThrIleLeuGlyLeuValMetArgPheThrArgLysGluLeuGluArg 405
 QY 1486 CCTTCAAGTGGCCTGTTTCATCCAGCTTTTGTTCCTTCACATGCTCTTCTATGCT 1545
 Db 406 ProIleLysValProValIleProValIleProValLeuMetThrLeuIleSerValPheLeuVal 425
 QY 1546 GGCCTTTCCTCTATTCGAGCCCATTTAGTACAGGATGCTGCTGCTGCTGCTGCTGCT 1602
 Db 426 LeuAlaProIleIleSerLysProThrTrpGluTyrLeuTyrCysValLeuPheIleLeu 445
 QY 1603 ACTGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1662
 Db 446 SerGlyLeuLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 463
 QY 1663 ATAATGTCAGAGAAATAACACAGAACATTAACAAATATATCTGGAAGTTGTACCAAGAA 1722
 Db 464 LysIleSerLysProIleThrMetHisLeuGlnMetLeuMetGluValValProGlu 483
 QY 1723 GAT 1725
 Db 484 Glu 484

RESULT 11
 BAT1_RAT
 ID BAT1_RAT STANDARD; PRT; 487 AA.
 AC P8252;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE B(0,+)-type amino acid transporter 1 (B(0,+)-AT) (Glycoprotein-
 associated amino acid transporter b0,+AT1).
 GN SLC7A9 OR BAT1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N. A., FUNCTION, TISSUE SPECIFICITY, SUBUNITS, AND
 RP SUBCELLULAR LOCATION.
 RC TISSUE=Kidney;
 RX MEDLINE=99436074; PubMed=10506124;
 RA Chairoungdua A., Segawa H., Kim J.Y., Miyamoto K.-I., Haga H.,
 RA Fukui Y., Mizoguchi K.-I., Ito H., Takeda E., Endou H., Kanai Y.;
 RT Identification of an amino acid transporter associated with the

RT cystinuria-related type II membrane glycoprotein.";

RL J. Biol. Chem. 274:28845-28848 (1999).

CC -1- FUNCTION: INVOLVED IN THE HIGH-AFFINITY, SODIUM-INDEPENDENT

CC TRANSPORT OF CYSTINE AND NEUTRAL AND DIBASIC AMINO ACIDS (SYSTEM

CC B(0,+)-LIKE ACTIVITY). THOUGHT TO BE RESPONSIBLE FOR THE HIGH-

CC AFFINITY REABSORPTION OF CYSTINE IN THE KIDNEY PROXIMAL TUBULE.

CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER WITH THE AMINO ACID

CC TRANSPORT PROTEIN SLC3A1.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

CC -1- TISSUE SPECIFICITY: KIDNEY AND SMALL INTESTINE. IN THE KIDNEY

CC LOCALIZED TO THE APICAL MEMBRANE OF THE PROXIMAL TUBULES.

CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID

CC PERMEASES. GLYCOPROTEIN-ASSOCIATED AMINO ACID TRANSPORTERS (GPAT)

CC SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB029559; BAA85186.1; .

DR InterPro; IPR002293; AA/rei_pmease1.

DR InterPro; IPR004841; Permease.

DR Pfam; PF00324; aa_permeases; 1.

DR Transport; Amino-acid transport; Transmembrane.

KW TRANSMEM 30 50 POTENTIAL.

FT TRANSMEM 64 84 POTENTIAL.

FT TRANSMEM 100 120 POTENTIAL.

FT TRANSMEM 179 199 POTENTIAL.

FT TRANSMEM 255 275 POTENTIAL.

FT TRANSMEM 297 317 POTENTIAL.

FT TRANSMEM 349 369 POTENTIAL.

FT TRANSMEM 375 395 POTENTIAL.

FT TRANSMEM 410 430 POTENTIAL.

FT TRANSMEM 435 455 POTENTIAL.

SQ SEQUENCE 487 AA; 53657 MW; EB43A1FC8BDC3A7 CRC64;

Alignment Scores:

Pred. No.: 8,9e-77 Length: 487

Score: 1074.50 Matches: 203

Percent Similarity: 66.52% Conservative: 103

Best Local Similarity: 44.13% Mismatches: 151

Query Match: 26.08% Indels: 3

DB: 1 Gaps: 2

US-09-667-170A-440 (1-2239) x BAT1_RAT (1-487)

QY 346 CTGAGAGGAAAGTCACTTACTGAGGGAGTCTCCATTATCATTTGGCACCATCATTTGGA 405

DB 26 LeuGlnLysGluValGlyLeuLeuSerGlyIleCysIleIleValGlyThrIleIleGly 45

QY 406 GCAGAAATCTTCTCTCTAAGGCGTGTCTCCAGAACACGGGAGCGTGGCGATGTCT 465

DB 46 SerGlyIlePheIleSerProLysSerValLeuAlaAsnThrGluSerValGlyProCys 65

QY 466 CTGACCATCTGGAGGTGTGTGGGTCCTGCTACTATTGAGGCTTGTGCTTATGCTGAA 525

DB 66 LeuIleIleThrAlaAlaCysGlyValLeuLeuAlaThrLeuGlyAlaLeuCysPheAlaGlu 85

QY 526 TTGGGAACAATATAAGAAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTGGT 585

DB 86 LeuGlyThrMetIleThrLysSerGlyGlyGlyProTyrProTyrLeuMetGluAlaPheGly 105

QY 586 CCATTACAGGTTTGTGAGGTCCTGAGTCTCTCATATACGCTTCGAGTCT 645

DB 106 ProIleProAlaTyrLeuPheSerTyrThrSerLeuIleValMetLysProSerSerPhe 125

QY 646 GCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTTATTTCAATGTGAA 705

DB 126 AlaIleIleCysLeuSerPheSerGluTyrValCysAlaAlaPheTyrLeuGlyCysArg 145

QY 706 ATCCCTGAACCTTGGCATCAAGCTCATTACAGCTGTGGGCATATAACTGTAGTATGGTCTTA 765

DB 146 ProProAlaValValLysLeuLeuAlaAlaAlaIleLeuLeuIleThrThrVal 165

QY 766 AATAGCATGAGTGTCTGAGTGGAGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTC 825

DB 166 AsnAlaLeuSerValArgLeuGlySerTyrValGlnAsnValPheThrAlaAlaLysLeu 185

QY 826 ACAGCAATCTGTATATTAATTAATCTGAGTATGAGCTAATTAAGGTCAACACGAG 885

DB 186 ValIleValAlaIleIleIleSerGlyLeuValLeuLeuAlaGlnGlyAsnValLys 205

QY 886 AACTTTAAAGACCCCTTTTCAGAGAGATTCAGATTTACCGGTTGCCACTGCTTTT 945

DB 206 AsnPheGlnAsnSerPheGluGlySerGlnThrSerValGlySerIleSerLeuAlaPhe 225

QY 946 TATTATGAATGATGATGATGCTGGTGGTGTACCTCACTTGTACTTGTACTGAAGAAGTA 1005

DB 226 TyrAsnGlyLeuTrpAlaTyrAspGlyTyrPAsnGlnLeuAsnIleThrGluGluLeu 245

QY 1006 GAAACACCTGAAAAACCATTTCCCTTGCATATATATATATATATATATATATATAT 1065

DB 246 ArgAsnProTyrArgAsnLeuProMetAlaIleValIleGlyLeuProLeuValThrVal 265

QY 1066 GGCTATGCTGTGACAAATGCGCTACTTTACGACCATTAATCTGAGGAGCTGTGCTT 1125

DB 266 CystyrIleLeuMetAsnIleAlaTyrPheThrValMetThrProThrGluLeuLeuGln 285

QY 1126 TCAATCGAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTT 1185

DB 286 SerGlnAlaValAlaValThrPheGlyAspArgValLeuTyrProAlaSerTrpValVal 305

QY 1186 CCGATCTTGTGCTCTCTCTGCTTGGTCCATGAACGCGTGTGTGTGTGTGTGTGTCTCC 1245

DB 306 ProLeuPheValAlaPheSerThrIleGlyAlaAlaAsnGlyThrCysPheThrAlaGly 325

QY 1246 AGGTTATTTCTATGTTGCTGCTGAGAGGTCACCTCCAGAAATCCCTCTCATGATTCAAT 1305

DB 326 ArgLeuIleTyrValAlaGlyArgGluGlyHisMetLeuLysValLeuSerTyrIleSer 345

QY 1306 GTCCGCAACACACCTCTCTACAGCTGTATTTGTTGACCCCTTTGACAAATGATGATG 1365

DB 346 ValLysArgLeuThrProAlaProAlaLeuValPheThrGlyIleIleAlaIleIleTyr 365

QY 1366 CTCCTCTCTGAGAGCTCGACAGCTTTTGAATTTCTCTCAGTTTTCAGGTCGCTTTT 1425

DB 366 IleIleProGlyAspIleAsnSerLeuValAsnTyrPheSerPheAlaAlaTrpLeuPhe 385

QY 1426 ATGGGCTGGCAGTGTGCTGCTGATTTATCTTCGATACAAATGCCAGATATGATCGT 1485

DB 386 TyrGlyMetThrIleLeuGlyLeuValValMetArgPheThrArgLysAspLeuGluArg 405

QY 1486 CCTTCAAGTCCAGTGTTCATCCAGCTTTGTTGTTTCTTCACATGCTCTTCATGTT 1545

DB 406 ProIleLysValProIlePheIleProIleIleValIleLeuValSerValPheLeuIle 425

QY 1546 GCCTTTCCCTCTATTCGAGCCCATTTAGTACAGGATTTGGCTTCGTCATC---ACTCTG 1602

DB 426 LeuAlaProIleIleSerSerProAlaTrpGluTyrLeuTyrCysValLeuPheIleLeu 445

QY 1603 ACTGGAGTCCCTCGCTATTATCTCTTTTATATATATATATATATATATATATATAT 1662

DB 446 SerGlyLeuIlePheTyrPheLeuPheValHisTyr-----LysPheArgTrpAlaGln 463

QY 1663 ATATGTACAGAAATAACACAGAACATTAACAAATAATTAATGAAAGTGTGTACAGAGAA 1722

DB 464 LysIleSerArgProIleThrThrLysHisLeuGlnMetLeuMetGluValValProGlu 483

RESULT 12

BAT1_MOUSE

ID BAT1_MOUSE

AC Q90XAG;

STANDARD; PRT; 487 AA.

Qy 1426 ATGGGCTGGCAGTGTGCTGGCTGATTATCTTCGATACAAATGCGCCAGATATGCATCGT 1485
 Db 386 TyrGlyMetThrIleLeuGlyLeuValMetArgPheThrArgLysAspLeuGluArg 405
 Qy 1486 CTTTCAAGTGGCAGCTGTTTCATCCAGCTTTGTTTCTTCCTTCATGCTCTTCATGGTT 1545
 Db 406 ProIleLysValProLeuPheIleProIleValIleLeuValSerLeuPheLeuIle 425
 Qy 1546 GCCCTTCCCTCTATTTGGGACCATTTAGTACAGGATTGGCTTCGTCATC---ACTCTG 1602
 Db 426 LeuAlaProIleLeuSerGluProAlaTrpGluTyrLeuTyrCysValLeuPheIleLeu 445
 Qy 1603 ACTGGAGTCCCTGCTGCTTATCTTCTTTATATATGGGACAAAGAACCCAGGTGGTTTGA 1662
 Db 446 SerGlyLeuIlePheTyrPheLeuPheValTyrTyr-----LysPheGlyTrpAlaGln 463
 Qy 1663 APAATGTCAGAGAAATAACAGCAACATTAACATAATATCTGCAAGTGTGTACCAAGAGAA 1722
 Db 464 ArgIleSerArgProValThrLysHisLeuGlnMetLeuMetGluValValProGlu 483

RESULT 13
 MUPL_YEAST
 ID MUPL_YEAST STANDARD; PRT; 574 AA.
 AC P50276;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE High-affinity methionine permease.
 GN MUPL OR YGR055W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RX MEDLINE=97049127; PubMed=8893857;
 RA Isnard A.D., Thomas D., Surdin-Kerjan Y.;
 RT "The study of methionine uptake in Saccharomyces cerevisiae reveals a
 RL new family of amino acid permeases";
 RL J. Mol. Biol. 262:473-484(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Entian K.D., Rose M., Koettler P., Roehmer A., Sehrsam I.,
 RA Hempel S.;
 RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: HIGH AFFINITY PERMEASE FOR METHIONINE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. (Potential).
 CC -!- SIMILARITY: TO YEAST LOW AFFINITY METHIONINE PERMEASE (MUP3).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U40316; AAB63529.1; -;
 CC SGD; Z72840; CNA97055.1; -;
 DR InterPro; IPR002293; AA/rel_permease1.
 DR InterPro; IPR004760; L_AA_transport.
 DR InterPro; IPR004841; Permease.
 DR Pfam; PF00324; aa_permeases; 1.
 DR TIGRFAMs; TIGR00911; 2A0308; 1.
 DR Transport; Amino-acid transport; Transmembrane.
 KW TRANSMEM 62 82 POTENTIAL.
 FT TRANSMEM 93 113 POTENTIAL.
 FT TRANSMEM 141 161 POTENTIAL.
 FT TRANSMEM 183 203 POTENTIAL.
 FT TRANSMEM 208 228 POTENTIAL.
 FT TRANSMEM 294 314 POTENTIAL.

FT TRANSMEM 341 361 POTENTIAL.
 FT TRANSMEM 419 439 POTENTIAL.
 FT TRANSMEM 436 476 POTENTIAL.
 FT TRANSMEM 491 511 POTENTIAL.
 SQ SEQUENCE 574 AA; 63221 MW; F95D342814710837 CRC64;
 Alignment Scores: Length: 574
 Pred. No.: 3 05e-21 Matches: 135
 Score: 368.50 Conservative: 104
 Percent Similarity: 45.61% Mismatches: 199
 Best Local Similarity: 25.76% Indels: 86
 Query Match: 8.94% Gaps: 22
 DB: 1
 US-09-667-170A-440 (1-2239) x MUPL_YEAST (1-574)
 Qy 280 GACGCCCTTTTCAGGAAGACGCGCTTTTCAGGAAGACGCGCTTTTCAGGAAGACGAGAA 339
 Db 36 AspAlaAspAsnGlyAlaSerAspPheGluAlaGlyGlnPheAla---ThrGluLeu 54
 Qy 340 CTGCAGCTGAAGAGAAAGTCACTTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATC 399
 Db 55 AspGlnGlyGluLysGlnLeuGlyIleLeuSerCysIleGlyLeuIleCysAsnArgMet 74
 Qy 400 ATGGAGCAGGAATCTTCATCTCTCCTAAGGGGTCTCCAGAACACGGCAGCGTGGGC 459
 Db 75 LeuGlyThrGlyValPheAlaValSerThrIleTyrThrLeuCysGlySerValGly 94
 Qy 460 ATGTCCTGACCATCTGGACGGTGTGGGTCTCTGACACTATTTGGAGCTTTGCTCTTAT 519
 Db 95 LeuAlaLeuIleMetTrpAlaValGlyAlaIleAlaIleSerGlyLeuTyrValTyr 114
 Qy 520 GCTGAATTTGGGAACAACATATAAGAAATCTCGAGGTCACTTACACATATATTTTGAAGTC 579
 Db 115 MetGluPheGlyThrAlaIleProLysAsnGlyGlyLysAsnTyrLeuGluAlaIle 134
 Qy 580 TTGTGTCATTCACGAGCTTTGTAGAGTCTGGGTGGAACCTCTCATATAACGCCCTGCA 639
 Db 135 PheArgLys---ProLysPhe----- 140
 Qy 640 GCTACTGCTGATATCCCTGCGCATTTGGAGCTACATCTTGGACCATTTTATTATCAA 699
 Db 141 -----PheIleThrCysMetTyrAlaAlaTyrIle-----PhePheLeuGly 154
 Qy 700 TGTGAATCCCT-----GAACCTTGGATCAAGCTTCATACAGCT----- 738
 Db 155 TrpAlaAlaGlyAsnSerIleAsnThrAlaIleMetPheLeuThrAlaAlaAspThrGlu 174
 Qy 739 -----GTGGCATAACTGTAGTG-----ATGTCCTA 765
 Db 175 ValThrLysTrpAsnGlnArgGlyIleGlyValAlaValPhePheAlaPheLeuIle 194
 Qy 766 AATGATGATGTCAGCTGAGCGCGCGGATCCAGATTTCTTAACTTTTGAAGCTC 825
 Db 195 AsnSerLeuAsnValLysIleGlyLeuTyrLeuGlnAsnIleLeuGlyIlePheLysIle 214
 Qy 826 ACAGCAATCTGATATATAGTCCCTGAGTATGACGCTAATTAAGGT----- 876
 Db 215 GlyIleValLeuPheIleSerIleThrGlyTrpValAlaLeuGlyGlyLeuLysAsp 234
 Qy 877 -----CAAACGCGAAGCTTTAAAGACGCTTTTCAGGAAGAGAGATTCAAGTATTACG--- 927
 Db 235 GlyTyrGlnSerHisAsnPheArgAsnAlaPheGluGlyThrGluThrAlaThrAlaTyr 254
 Qy 928 CGGTGGCCTGCTGCTTTTATTATGAATGTATGATGATGCTGGCGGTGTTTACCTCAAC 987
 Db 255 GlyIleValAsnAlaLeuTyrSerValIleTrpSerPheValGlyTyrSerAsnValAsn 274
 Qy 988 TTTGTTTACTGAAGTAGAAGAACCCCTGAAAAACCATTCCTTCGATATATGATATCC 1047
 Db 275 TyrAlaLeuGlyGluValLysAsnProValArgThrLeuLysIleAlaGlyProThrSer 294
 Qy 1048 ATGGCCATTGTGCACCATTTGGCTATGTGCTGACAAATGTGGCCTTACCTTACGACCATTAT 1107

Db 295 MetValPheLeuAlaIleIleTyrIlePheValAlaIleAlaIleValValPro 314
QY 1108 GCTGAGCAGCTGCTTCTTCAATGAGTG---GCAGTGACCTTTTCTGAGCGGCTACTG 1164
Db 315 LysAspLysLeuIleSerLysLeuIleLeuAlaAlaAspPheAspIleValPhe 334
QY 1165 ---GGAAATTTTCATTAGCAGTTCGATCTTTGTGCTCTCTGCTTGGCTCCATG 1221
Db 335 GlyGlyGlnAlaLysArgAlaAlaAlaLeuValGlyLeuSerAlaLeuGlyAsnVal 354
QY 1222 AAGCGTGCTGTTCTGCTCCAGGTATTTCTATGTGTGCTCCGAGAGGGTCACTT 1281
Db 355 LeuSerValIlePheSerGlnGlyArgIleGlnLeuLeuGlyArgGlyValLeu 374
QY 1282 CCA---GAAATCTCTCCATGATTCATGTCGCGAAGCACACATCTCTCTA 1326
Db 375 ProPheSerAsnPheAla-SerSerLysProPheAsnSerProMetValGlyLeuPh 394
QY 1327 CAGCTGTTATGTTTGGACCTTTGACATGATGATGATGATGATGATGATGATGAT 1386
Db 394 eGlnHisPheIleValCysThr--ValThrIleLeuAlaProProGlyAspAlaTyr 413
QY 1387 ACTCTTTTG---AATTTCTCTAGTTTTCGAGGTGGCTTTTATGCTGGCTGGCAGTTGCT 1443
Db 414 LeuLeuValGlnAsnLeuIleSerTyrProMetAsnIleIleAsnPheAlaIleSerAla 433
QY 1444 GCGCTGATTTATCTT---CGATACAAATGCCAGATATGATGATGATGATGATGAT 1491
Db 434 GlyLeuLeuTyrIleTyrTrpGlnArgGlnGlyLysIleGluTrpAsnProIle 453
QY 1492 AAGGTGCCCATGTTCCATCCAGCTTTGTTTCTTCACATGCTCTCTTCATGTTGCCCTT 1551
Db 454 LysAlaGlyValPheValThrGlyPhePheThrLeuSerAsnLeuIleIleAla 473
QY 1552 -----TCCCTCTATTCGGAC---CCATTT----- 1572
Db 474 ProTyrValProProSerAsnGlyLeuSerValTyrSerSerMetProTyrTrpIleHis 493
QY 1573 -----AGTACAGGATGCTGCTGCTCATCTCTGAGTCCCTGCTATTAT 1623
Db 494 CysValIleAlaTrpGlyIlePhePhe-----PheGlyGlyVal-----TyrTyr 508
QY 1624 CTCTTTATTATGGACAAG---AAACCCAGGTGG-----TTTGAATATATGTCAGAG 1674
Db 509 -----ValValTrpAlaGlnLeuLeuProArgTrpGlyHisTyrLysLeuValSerLys 526
QY 1675 AAAATA 1680
Db 527 AspVal 528

RESULT 14
MUP3_YEAST
ID MUP3_YEAST STANDARD; PRT; 546 AA.
AC P38734;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
GN Low-affinity methionine permease.
OS MUP3 OR YHL036W.
OC Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,

Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.,
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII.";
RL Science 265:2077-2082(1994).
[2]
RP CHARACTERIZATION.
RX MEDLINE=97049127; PubMed=8893857;
RA Isnard A.D., Thomas D., Surdin-Kerjan Y.;
RT "The study of methionine uptake in Saccharomyces cerevisiae reveals a
new family of amino acid permeases.";
RL J. Mol. Biol. 262:473-484(1996).
CC -!- FUNCTION: VERY LOW AFFINITY PERMEASE FOR METHIONINE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -!- SIMILARITY: TO YEAST HIGH AFFINITY METHIONINE PERMEASE (MUP1).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL: U11583; AB65048.1; -
DR PIR; S48932; S48932.
DR SGO; S0001028; MUP3.
DR InterPro; IPR002293; AA/rel_permease.
KW Transport; Amino-acid transport; Transmembrane.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 320 340 POTENTIAL.
FT TRANSMEM 347 367 POTENTIAL.
FT TRANSMEM 394 414 POTENTIAL.
FT TRANSMEM 424 444 POTENTIAL.
FT TRANSMEM 460 480 POTENTIAL.
FT TRANSMEM 495 515 POTENTIAL.
SQ SEQUENCE 546 AA; 60619 MW; 9C9767A181FDDCAE CRC64;

Alignment Scores:
Pred. No.: 2.8e-19 Length: 546
Score: 343.50 Matches: 138
Percent Similarity: 42.96% Conservative: 100
Best Local Similarity: 24.91% Mismatches: 213
Query Match: 8.34% Indels: 103
DB: 1 Gaps: 25

US-09-667-170A-440 (1-2239) x MUP3_YEAST (1-546)
QY 247 CTTTCCTGGGCAAC-----AAGGAGCCACCTGGGAGGCGCCTTT----- 288
Db 3 ProLeuLeuPheAsnSerGlyLysAlaAsnProSerGlnAspValPheIleAspValGlu 22
QY 289 -----TCAGGAAGAGAGCGCTTTTCAGGAAGAGAG 318
Db 23 ValGlyAspIleThrTyrTyrGlySerThrAsnThrGlySerPheSerMetAsp 42
QY 319 GCCTTTTCAGGAAGAGAG---AAAGTCAGCTGAGAGGAAGTCACTTTTACTGAGGGGA 375
Db 43 ThrValGluAlaGlnAlaIleLysAlaGluThrAlaArgPheMetGluValProGlnGly 62
QY 376 -----GTCTCCATTATTCATGTCACCATTCATTGGAGCAGGA 411
Db 63 ArgHisLeuGlyValPheSerThrValValLeuPheValSerArgIleMetGlySerGly 82
QY 412 ATCTTCATCTCTCTTAAGGGCTGCTCCAGAACACGGGCGAGCTGTCTCTG--- 468
Db 83 IlePheAlaValProSerValIleLeuLeuAsnThrGly-----GlyAsnLysLeuIle 100

1465 AAATGCCAGATATGATCGTCTCTTCAAGGTGCGCATGTTTCATCCAGCTTTGTTTCC 1524
 144 Lys -----HisLysAsn-----GluValProGluIleArgAlaSer-----Thr 460
 1525 TTCCATGCTCTTTCATGCTTCCCTTCCCTCTCTAT-----TCG 1563
 461 PheGlyValLeuAlaIleIleThrLeuSerLeuTyrMetLeuMetAlaProPhePheAla 480
 1564 GACCCATTAGTACAGGGATTCGCTTCGTCATCTCTGACTGGAGTCCCTGCTATATAT 1623
 481 AspProSerLeuAsnArgValGlyPhe-----LeuProProTyrGln 494
 1624 CTC -----TTTATTATATGGACAAGAAA----- 1647
 495 IleMetSerLeuValIleValAlaCysPhePhePheTrpLeuValIlePheValLeu 514
 1648 ---CCCAGGTGGTTTAGAATA---ATCTCAGAGAAAATAACC 1683
 515 LeuProLysPhePheHisTyrIleValLeuLeuProLysIleThr 528

RESULT 15
 YHFM_ECOLI STANDARD; PRT; 445 AA.
 ID YHFM_ECOLI STANDARD; PRT; 445 AA.
 AC P45539; P76686;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical transport protein yhfM.
 GN YHF OR B3370.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R.; Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -!- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
 CC PERMEASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U18997; AAA58167.1; ALT_SEQ.
 CC EMBL: A5000413; AAC76395.1; ALT_INIT.
 CC EcoGene: EGI2908; yhfM.
 CC InterPro: IPR002293; AA/rel.prmase1.
 CC InterPro: IPR004841; Permease.
 CC Pfam: PF00324; aa-permeases; 1.
 CC Hypothetical protein; Transport; Transmembrane; Inner membrane;
 CC Complete proteome.
 CC TRANSMEM 10 30 POTENTIAL.
 CC TRANSMEM 38 58 POTENTIAL.
 CC TRANSMEM 93 113 POTENTIAL.
 CC TRANSMEM 121 141 POTENTIAL.
 CC TRANSMEM 155 175 POTENTIAL.
 CC TRANSMEM 181 201 POTENTIAL.
 CC TRANSMEM 236 256 POTENTIAL.
 CC TRANSMEM 273 293 POTENTIAL.
 CC TRANSMEM 334 354 POTENTIAL.

469 -----ACCATCGGAGGTGCTGGGGTCTGTCACATATTTGGAGCTTTGTCTTATGCT 522
 101 TyrPheAlaIleTrpValPheSerAlaIleAlaPheAlaGlyLeuTyrLeuPheLeu 120
 523 GAATTGGGAACAACATAAAGAAATCTGGAGGTCAFTACACATATATTTTGGAACTCTTT 582
 121 GluPheGlySerTrpIleProLysSerGlyGlyArgLysAsnPheLeuGluArgSerPhe 140
 583 GGTCCATTACAGCTTTTGTACAGAGTCTGGTGGAACTCTCATATACATCCCTGCAGCT 642
 141 GluArgProArgLeuLeuIleSerValValPheSerCysTyrSerValLeuThrGlyTyr 160
 643 ACTGCTGTGATATCCCTGGCATCTGGACGCTACATCTCGAACCATTTTATTCATGCT 702
 161 AlaLeuThrGlySerIleValPheGlyLysTyrValLeuSerAlaPheGlyVal ----- 178
 703 GAAATCCCTGAACCTGGCATCAAGCTCATACAGCTGTGGGCATACTAGTAGTGATGTC 762
 179 --ThrAspAspSerTrpSerLysTyrValSerIleSerPheIleIlePheAlaValLeu 197
 763 CTAATACATGAGTGCAGCTGCGGCGCGGATCCAGATTTCTTAACCTTTTGCAG 822
 198 IleHisGlyValSerValArgHisGlyValPheIleGlnAsnAlaLeuGlyGlyLeuLys 217
 823 CTCACAGCAATCTGATATATAGTCCCTGGAGTATGACGCTAAT----- 870
 218 LeuIleMetIleValLeuMetCysPheAlaGlyLeuTyrThrLeuPheTyrLysSer 237
 871 AAGGTCAA-----ACGCAGAACTTTAAAGACGCTTTTCAGGA 909
 238 ThrGlyGlnValAlaTrpAspLeuProValThrGlnValGluLysAspSerLeu----- 255
 910 AGAGATTCAAGTATACCGGTGCGCTGCTTTTATTGAAATGATGATGATGCT 969
 256 -----LeuSerValSerSerIleAlaThrAlaPheIleSerPheCysPheSer 273
 970 GGTGCTTTTACTCAACTTCTTACTGAAGATGAGAAACCTCGTAAACCAATCCCT 1029
 274 GlyTrpAspThrValHisThrValThrSerGluIleLysAsnProValLysThrLeuLys 293
 1030 CTGCAATATGATATCCATGCGCATGTCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1089
 294 ValSerGlyProLeuSerLeuIleLysCysPheValCysTyrThrMetMetValAla 313
 1090 TACTTTACGACCAATTAATGCTGAGGAGCTG-----CTTTCAATGCGAGT 1137
 314 TyrLeuLysValLeuThrTyrGluGluIleValSerAlaGlyProLeuValGlySerVal 333
 1138 GCAGTGACC-----TTTCTGAGCGCTACTGGGAAATTTCTCATTAGCAGTCCGATC 1191
 334 LeuPheThrLysLeuPheGlyProArgValGlyGlyLysPhe-----IleAlaPhe 350
 1192 TTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1251
 351 SerIleAlaIleSerAlaAlaSerAsnIleLeuValValIleTyrSerIleSerArgVal 370
 1252 TTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1311
 371 AsnGlnGluIlePheLysGluGlyTyrLeuPro-----PheSerIleHisMetSer 387
 1312 AAGCAC-----ACTCCTTACAGCTGTTATGTTTGTGACCTTTGACCAATG 1359
 388 LysAsnTrpProPheAspAlaProLeuProSerIleSerLeuCysGlyPheIleThrIle 407
 1360 -----ATAAGTCTCTCTGGAGAGCTCGACAGT-----CTTTGATATTTCTTC 1404
 408 AlaTrpIleIleLeuProLysGluGlySerPheAsnTyrLeuValSerMetAsp 427
 1405 AGTTTTCGAGGTGGCTTTTATTGGCTGGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1464
 428 GlyTyrGlyAsnGlnPhePheLeuLeuValAlaIleGlyLeuPheIleTrpArgPhe 447

Search completed: April 16, 2003, 16:10:37
Job time : 41 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 15:45:20 ; Search time 96.5 Seconds
(without alignments)
9561.440 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 4120
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cun2_1/USPTO_spool/US09667170/runat_04042003_090917_20296/app_query.fasta_1.2375
-DB=SPTREMBL_21 -QFMR=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09667170_ECGN_1_122_runat_04042003_090917_20296 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	2424	58.8	494	4	Q9BYH2	Q9byh2 homo sapien

ID	Q9BYH2	PRELIMINARY;	PRT;	494 AA.
AC	Q9BYH2;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Cystine/glutamate exchanger.			
GN	HACT.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Kim J.Y., Chairoungdua A., Cha S.H., Segawa H., Matsuo H., Kim D.K.,			
RA	Endou H., Kanai Y.;			
RT	"Human cystine/glutamate exchanger: cDNA cloning and upregulation by			
RT	oxidative stress in glioma cells.;"			
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB040875; BAB40574.1;			
DR	InterPro; IPR002293; AA/rel_pmaseal.			
DR	InterPro; IPR004760; LAA_transport.			
DR	InterPro; IPR004841; Permease.			
DR	Pfam; PF00324; aa_permeases; 1.			

ALIGNMENTS

RESULT 1

ID	Q9BYH2	PRELIMINARY;	PRT;	494 AA.
AC	Q9BYH2;			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Cystine/glutamate exchanger.			
GN	HACT.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Kim J.Y., Chairoungdua A., Cha S.H., Segawa H., Matsuo H., Kim D.K.,			
RA	Endou H., Kanai Y.;			
RT	"Human cystine/glutamate exchanger: cDNA cloning and upregulation by			
RT	oxidative stress in glioma cells.;"			
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB040875; BAB40574.1;			
DR	InterPro; IPR002293; AA/rel_pmaseal.			
DR	InterPro; IPR004760; LAA_transport.			
DR	InterPro; IPR004841; Permease.			
DR	Pfam; PF00324; aa_permeases; 1.			

DR TIGRFAMS: TIGR00911; 2A0308; 1. BSC57A7EA4DDB0B1 CRC64;
SQ SEQUENCE 494 AA; 54423 MW; 505 AA; 55110 MW; D0E6ECC11B8594E1 CRC64;

Alignment Scores:
Pred. No.: 8, 38e-205 Length: 494
Score: 2424.00 Matches: 478
Percent Similarity: 96.37% Conservative: 0
Best Local Similarity: 96.37% Mismatches: 0
Query Match: 58.83% Indels: 18
DB: 4 Gaps: 1

US-09-667-170A-440 (1-2239) x Q9BYH2 (1-494)

QY 184 AAGCCTGTTGTGTCACCATCTCCAAAGGAGGTTACCTGCGAGGAAATGTTAAACGGGAGG 243
Db 4 LysProValValSerThrIleSerLysGlyGlyTyrLeuGlnGlnValAsnGlyArg 23

QY 244 CTGCTTCCCTGGGCAACAGGAGGACCTGGGAGGAGGCGCTTTTCAGGAAGAGACGCC 303
Db 24 LeuProSerLeuGlyAsnLysGluProProGlyGln----- 35

QY 304 TTTTCAGGAAGAGCGCTTTTCAGGAAGAGAGAAAGTCCAGCTGAAGAGAAAGTCACT 363
Db 36 -----GluLysValcLlnLeuLysArgLysValThr 45

QY 364 TTACTAGGGGAGTCCCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 423
Db 46 LeuLeuArgGlyValSerIleIleGlyThrIleIleGlyValAlaGlyIlePheIleSer 65

QY 424 CCTAAGGCGTGTCCAGAACACGGGAGGCGTGGGATGTCCTGACCATCTGGAGGTTG 483
Db 66 ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleTrpThrVal 85

QY 484 TGTGGGTCCTGTCACATATTGGAGCTTCTTATTCGTGAATTTGGGACACACTATAAG 543
Db 86 CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrThrIleLys 105

QY 544 AAATCTGAGGTCATTACACATATTTTGAAGTCTTTGGTCCATTACACAGCTTTTGA 603
Db 106 LysSerGlyGlyHisTyrThrTyrIleLeuGluValPheGlyProLeuProAlaPheVal 125

QY 604 CAGTCTGGGTGGAACTCCTCATATACGGCTGACGCTACTGCTGTGATATCCCTGGCA 663
Db 126 ArgValTrpValGluLeuLeuIleArgProAlaAlaThrAlaValIleSerLeuAla 145

QY 664 TTTGAGCGTACATCTCTGGACCATTTTATTCATTAATGTGAATCCCTGAACTTGGATC 723
Db 146 PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle 165

QY 724 AAGCTCATTTACAGCTGTGGGCATACTAGTGTGATGTCCTTAATAGCATGAGTGCAGC 783
Db 166 LysLeuIleThrAlaValAlaGlyIleThrValValMetValLeuAsnSerMetSerValSer 185

QY 784 TGGAGCGCCGGATCCAGATTTCTTAACCTTTTTCAGAGCTCAGCAATCTCTGAATTT 843
Db 186 TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIle 205

QY 844 ATAGTCCCTGGAGTATGACGCTAATTAAGCTCAACGCGAGAACTTTAAAGAGCGCTTT 903
Db 206 IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPheLysAspAlaPhe 225

QY 904 TCAGGAAGAGATTCAGATATTACGGGTTGGCAGCTGGCTTTTATTTATGATGATGATGCA 963
Db 226 SerGlyArgAspSerSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla 245

QY 964 TATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAGTAGAAACCCCTGAAACAAAC 1023
Db 246 TyrAlaGlyTrpPheTyrLeuAsnPheValThrGluValGluAsnProGluLysThr 265

QY 1024 ATTCCCTTCAATATGATATCCATGCGCATTTGTCACCATGCTGCTGACAAAT 1083
Db 266 IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn 285

QY 1084 GTGGCCTACTTTACACACATTAATGCTGAGAGCTGCTGCTTTCAATGAGTGGCAGTG 1143
Db 286 ValAlaTyrPheThrThrIleAsnAlaGluLeuLeuLeuSerAsnAlaValAlaVal 305

QY 1144 ACCTTTTCTGAGCGCTACTCGGAAATTTCTCATTAGCAGTTCGATCTTTGTTGCCCTC 1203
Db 306 ThrPheSerGluArgLeuLeuGlyAsnPheSerLeuAlaValProIlePheValAlaLeu 325

QY 1204 TCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTTATTTCTATGTTGCG 1263
Db 326 SerCysPheGlySerMetAsnGlyGlyValPheAlaValSerArgLeuPheTyrValAla 345

QY 1264 TCTGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCGTCAGGACACACCTCT 1323
Db 346 SerArgGluGlyHisLeuProGluIleLeuSerMetIleHisValArgLysHisThrPro 365

QY 1324 CTACAGCTGTTATGTTTTCACCCCTTTTGACAAATGATAATGCTCTCTCTGGAGACCTC 1383
Db 366 LeuProAlaValIleValLeuHisProLeuThrMetIleMetLeuPheSerGlyAspLeu 385

QY 1384 GACAGTCTTTTGAATTTCTCAGTTTTCAGGTTGGCTTTTATTTGGCTGGCAGTTGCT 1443
Db 386 AspSerLeuLeuAsnPheLeuSerPheAlaArgTrpLeuPheIleGlyLeuAlaValAla 405

QY 1444 GGCTGATTTATCTCGATACAAATCCCGAGATATGATGCTCTTCAAGTGCCACTG 1503
Db 406 GlyLeuIleTyrLeuArgTyrLysCysProAspMetHisArgProPheLysValProLeu 425

QY 1504 TTTATCCAGCTTTGTTTTCCTTCATGATGCTCTTTCATGTTGGCTTTCCTCTATTTCG 1563
Db 426 PheIleProAlaLeuPheSerPheThrCysLeuPheMetValAlaLeuSerLeuTyrSer 445

QY 1564 GACCAATTTAGTACAGGATTTGCTTCATGATGCTCTTTCATGTTGGCTTTCCTCTATTTCG 1623
Db 446 AspProPheSerThrGlyIleGlyPheValIleThrLeuThrGlyValProAlaTyrTyr 465

QY 1624 CTCTTTATTTATGAGCAAGAACCCAGGTGGTGTGTAGATAATATGTC 1671
Db 466 LeuPheIleIleThrPheLysLysProArgTrpPheArgIleMetSer 481

RESULT 2
Q9TU26 PRELIMINARY; PRT; 505 AA.
ID Q9TU26
AC Q9TU26; 2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Blood-brain barrier large neutral amino acid transporter.
GN LAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=BRIN MICROVASCULATURE;
RX MEDLINE=98449811; PubMed=10518579;
RA Bodo R.J., Li J.Y., Nagaya M., Zhang C., Pardridge W.M.;
RT "Selective expression of the large neutral amino acid transporter at
the blood-brain barrier."
RL proc. Natl. Acad. Sci. U.S.A. 96:12079-12084 (1999).
DR EMBL; AF174615; AAF03694.1; .
DR InterPro; IPR002293; AA/reel_primease1.
DR InterPro; IPR004760; LAA.transport.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMS; TIGR00911; 2A0308; 1.
SQ SEQUENCE 505 AA; 55110 MW; D0E6ECC11B8594E1 CRC64;

Alignment Scores:
Pred. No.: 8, 06e-99 Length: 505
Score: 1221.00 Matches: 234

ПS-09-667-170A-440 (1-2239) x 057439 (1-507)

QY	256	GGCAACAAGGACCCACCTGGCGACGAGCGCCCTTTTCAGAGAGACGCGCTTTTCAGGAAGA	315
Db	32	GIYASnAlaGluProLysSerGlyAspGlyAlaAla-----	43
QY	316	GACGCGCTTTTCAGGAGAGAGAAAGTGCAGCTGAAGAGGAAACTCACTTTACTAGGGGA	375
Db	44	-----ValGluLeuGlnArgThrIleThrLeuValAsnGly	55
QY	376	GTCTCCATTATCATTTGGCACCACCATCATTTGGAGCAGGAATCTTCATCTCTCCTTAAGCGCTG	435
Db	56	ValAlaIleIleValGlyThrIleIleGlySerGlyIlePheValThrProThrGlyVal	75
QY	436	CTCCAGAACACGGCGAGCTGGCGATGCTCTGACCATCTGGACGGTGTGTGGGTCTCGT	495
Db	76	LeuArgGluAlaGlySerProGlyLeuSerLeuLeuValTrpAlaValCysGlyLeuPhe	95
QY	496	TCACTATTTGGAGCTTTGTCTTATCTCAATTTGGGAACAACATAAAGAAATCTGGAGGT	555
Db	96	SerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleSerLysSerGlyGly	115
QY	556	CATTACACATATTTTGGAGCTTTTGGTCCATTACCAGCTTTTTCAGAGTCTCGGTG	615
Db	116	AspTyrAlaTyrValLeuGluValTyrGlyAlaLeuProAlaPheLeuLysLeuTrpVal	135
QY	616	GAATCCTCATATAGCCCTCGACTACTGCTGTGATATCCCTGGCATTTTGGACGCTAC	675
Db	136	GluLeuLeuIleIleArgProSerGlnTyrIleValAlaLeuValPheAlaThrTyr	155
QY	676	ATTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTTGGATTCAGCTCAAGCTCATTACA	735
Db	156	LeuLeuLysProValPheProThrCysProValProAspAspAlaAlaLysIleValAla	175
QY	736	GCTGTGGGCATAACTGTAGTGTGTCTTAATACGATGAGTGTACAGTGGAGCGCCGG	795
Db	176	CysLeuCysIleLeuLeuLeuThrAlaIleAsnCysTyrSerValLysAlaAlaThrArg	195
QY	796	ATCCAGATTTTCTTAACCTTTTTCGACGCTCAGACAAATTCCTGATAATATAGTCCCTGGA	855
Db	196	ValGlnAspArgPheAlaAlaLysLeuLeuAlaLeuLeuLeuIleIleLeuGly	215
QY	856	GTATTGCAGCTAATTAAAGTCACACGAGAACTTTAAAGAC-----GCCTTTTCAGGA	909
Db	216	PheValGlnLeuGlyLysGlyValGluAspLeuLysProGluArgSerPheGluGly	235
QY	910	ACAGATTCAAGTATTACGCGGTGGCCACTGGCTTTTATTATGGAATCTATGCAATATGCT	969
Db	236	ThrSerThrAsnValGlyGlnTrpValLeuAlaLeuTyrSerGlyLeuPheAlaTyrGly	255
QY	970	GGCTGGTTTTACCTCAACTTTGTTACTGAAGAGTAGAACCCTGGAACCAACCATTTCCC	1029
Db	256	GlyTrpAsnTyrLeuAsnPheValValGluGluMetIleGluProTyrLysAsnLeuPro	275
QY	1030	CTTGCAAATATGATATCATCGCCATTCACCATTTGCATCGCTATGCTGCACAAATATGGCC	1089
Db	276	ArgAlaIleIleSerMetProIleValThrLeuValTyrValLeuThrAsnLeuAla	295
QY	1090	TACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCCAATTCAGTGCAGTGCAGTGCAC	1149
Db	296	TyrPheThrThrLeuThrProGluGlnMetLeuAsnSerGluAlaValAlaValAspPhe	315
QY	1150	TCGTAGCGGCTACTCGGGAATTTCTCATTAGCAGTTCGCGATCTTTGTGGCTCTCCCTG	1209
Db	316	GlyAsnTyrHisLeuGlyValAlaMetAlaTrpIleIleProValPheValGlyLeuSerCys	335
QY	1210	TTTGGCTCCATCAAGCGTGTGTGCTGTCTCCAGGTTATCTCATGTTCCGCTCTCGA	1269
Db	336	PheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPhePheValGlyAlaArg	355
QY	1270	GAGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGGCAGCACACTCCTCTACCA	1329
Db	356	GluGlyHisLeuProSerLeuLeuAlaMetIleHisProArgLeuLeuThrPrometPro	375

QY	1330	GCTGTTATTGTTTGCACCCCTTTGACAAAGATAATGCTCTCTCTGGAGACCTCGACAGT	1386
Db	376	SerLeuIlePheThrCysAlaMetThrLeuLeuIleAlaPheSerAspAlaPheSer	395
QY	1390	CTTTTGAATTCTTCAGTTTCCAGGTGGCTTTTATTGGCTGGCAGGTGCTGGGCTG	1449
Db	396	ValIleAsnPhePheSerPhePheAsnTrpLeuCysValAlaLeuAlaIleGlyMet	415
QY	1450	ATTATCTTCGATACAAATGCCAGATATGCATCTCTTTCAGGTGCCCATCTTCATC	1509
Db	416	MetTrpIleuArgTrpIleLysProGluLeuGluArgProIleLysValAsnIleLeuLeu	435
QY	1510	CCAGCTTTGTTTCTTCATCGCTCTTCATGGTGGCCCTTTCCTCTANTCGGCCCA	1569
Db	436	ProIlePhePheIleLeuAlaCysIlePheLeuIleValIleValSerPheTrpMetThrPro	455
QY	1570	TTTAGTACAGGGATTGGCTTCGTATCACTCTGACTGGAGTCCCGCTATTATCTCTTT	1629
Db	456	ValGluCysGlyIleGlyPheIleIleLeuThrGlyValProValTrpPhePheCly	475
QY	1630	ATTATATGGGACAGAAGAACCCAGGTGGTTTAGAATAATCTCAGAGAAAATAACCAAGACA	1689
Db	476	ValTrpTrpGlnAsnLysProAspTrpIleLeuHisGlyIleHisSerSerThrAlaLeu	495
QY	1690	TTACAAATAATCTGGAAGTTGTACAGAGAAGAA	1722
Db	496	LeuGlnIleValMetGluAlaValProGlnGlu	506

RESULT 4

AC	013020	PRELIMINARY;	PRT;	507 AA.
AD	013020;			
AC	013020;			
DT	01-JUL-1997	(TREMBLrel. 04, Created)		
DT	01-AUG-1998	(TREMBLrel. 07, Last sequence update)		
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)		
DT		Amino acid transporter chain (AmAt-L-1c) ASUR4.		
DE	ASUR4.			
GN	Xenopus laevis	(African clawed frog).		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
[1]				
RP	SEQUENCE FROM N.A.			
RP	TISSUE=KIDNEY;			
RC	MEDLINE=97324256; PubMed=9178633;			
RX	Spindler B., Mastroberardino L., Custer M., Verrey F.;			
RA	"Characterization of early aldosterone-induced RNAs in A6 kidney			
RT	epithelia.";			
RL	Plügers Arch. 434:323-331(1997).			
[2]				
RP	SEQUENCE FROM N.A.			
RP	TISSUE=KIDNEY;			
RC	MEDLINE=98421678; PubMed=9751058;			
RX	Mastroberardino L., Spindler B., Pfeiffer R., Skelly P.J., Loffli			
RA	Shoemaker C.B., Verrey F.;			
RT	"Amino acid transport by heterodimers of 4F2hc/CD98 and members			
RT	new permease family.";			
RL	Nature 395:288-291(1998).			
RL	EMBL: Y12716; CAA73254.1;			
DR	InterPro; IPR002293; AA/rel.prmeasel.			
DR	InterPro; IPR004760; L_AA_transp.			
DR	InterPro; IPR004841; Permease.			
DR	Pfam; PF00324; aa.permeases; 1.			
DR	TIGRfams; TIGR00911; 2A0308; 1.			
DR	CONFLICT 42 42 G -> A (IN REF. 0).			
FT	CONFLICT 199 199 A -> R (IN REF. 0).			
FT	CONFLICT 209 209 F -> L (IN REF. 0).			
FT	CONFLICT 391 391 N -> D (IN REF. 0).			
SO	SEQUENCE 507 AA; 55529 MW; 2A6741CE8AB721D CRC64;			

Alignment Scores:

Db	51	IlEthrLeuLeuAsnGlyValAlaIleValGlyThrIleIleGlySerGlyIlePhe	70
QY	418	ANCTCTCTAAGGGCGTCTCCAGAACACGGCAGCGTGGCATGCTCTGCACCACTGG	477
Db	71	ValThrProThrGlyLeuValGlySerProGlyLeuSerLeuValValTrp	90
QY	478	ACGGTGTGTGGGGTCCCTCTCACTATTTTGGAGCTTTGTCTTATGCTGAATGGGACAACT	537
Db	91	AlaValCysGlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThr	110
QY	538	ATPAAGAAATCTGGAGGTCATTACACATATATTTTGAAGCTCTTTGGTCCATTACCACT	597
Db	111	IleSerLysSerGlyLysPtyrAlaTyrMetLeuGluValTyrGlySerLeuProAla	130
QY	598	TTTGTACGAGTCTGGTGGAACTCCTCATATAGCCCTCAGCTACTGCTGTGATATCC	657
Db	131	PheLeuLysLeuTrpIleGluLeuLeuIleIleArgProSerGlnTyrIleValAla	150
QY	658	CTGGCATTTTGGAGCTACATCTCGAACCATTTTTATTCAAGTGAATCCCTGAACCT	717
Db	151	LeuValPheAlaThrTyrLeuLeuLysProValPheProThrCysProValProGluGlu	170
QY	718	GGCATCAAGCTCATACAGCTGTGGGCATACTCTAGTGTGTCTCTTAATAGCAGAT	777
Db	171	AlaAlaLysLeuValAlaCysLeuLysValLeuLeuLeuThrAlaValAsnCysTyrSer	190
QY	778	GTCAGCTGGAGCCCGGATCCAGATTTTCTTAACCTTTTGCAGACTCAGACAAATCTG	837
Db	191	ValLysAlaAlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAla	210
QY	838	ATAATATAGTCCCTGGAGCTATGACAGCTAATATAA-----GGTCAACGCCAG	885
Db	211	LeuIleIleLeuLeuGlyPheIleGlnMetGlyLysAspMetGlyGlnGlyAspAlaSer	230
QY	886	AACTTT-----AAGACGCCCTTTTCAGAGAGATTCACGATATTCGCGTTGCCACTG	939
Db	231	AsnLeuGlnGlnLysLeuSerPheGluGlyThrAsnLeuAspValGlyAsnIleValLeu	250
QY	940	GCTTTTATTATGAATGATGCATATGCTGGCTGGTTTACCTCAACTTTGTACTGAA	999
Db	251	AlaLeuTyrSerGlyLeuPheAlaTyrGlyTyrPasnTyrLeuAsnPheValThrGlu	270
QY	1000	GAGTAGAAACCTCGAAACCACTTCCCTTGGCAATATGTATATCCATGGCCATGTC	1059
Db	271	GluMetIleAsnProTyrArgAsnLeuProLeuAlaIleIleLeuSerLeuProIleVal	290
QY	1060	ACCATGGCATGTGCTGACAAATGGCTACTTTACGACCAATATGCTGAGAGCTG	1119
Db	291	ThrLeuValTyrValLeuThrAsnLeuAlaTyrPheThrThrLeuSerThrAsnGlnMet	310
QY	1120	CTGCTTTCAATGAGTGGCAGTGACCTTTTCTGACGGCTACTCGGAAATTTCTCATTA	1179
Db	311	LeuThrSerGluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSerTrp	330
QY	1180	GCAGTTCGATCTTTGTTCCTCTCCTTGGCTCCATGAACGGGTGTGTTTGGCT	1239
Db	331	IleIleProValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThr	350
QY	1240	GTCTCCAGGTTATCTATGTTGGTCTCGAGAGGTCACCTTCCAGAAATCCTCCCATG	1299
Db	351	SerSerArgLeuPhePheValGlySerArgGluGlyHisLeuProSerValLeuSerMet	370
QY	1300	ATTCATGTCCGAGACACACTCCTCTACAGCTGTTATTGTTTTCACCCCTTGACATG	1359
Db	371	IleHisProGlnLeuLeuThrProValProSerLeuValPheThrCysIleMetThrLeu	390
QY	1360	ATAATGCTCTTCTGGAGACCTCGACAGCTTTTGAATTTCTCAGTTTGGCCAGGTGG	1419
Db	391	MetTyrAlaPheSerArgAspIlePheSerIleIleAsnPhePheSerPhePheAsnTrp	410
QY	1420	CTTTTATTGGCTGGCAGTTGTGGCTGATTATCTTCGTATACAAATGCCAGATATG	1479
Db	411	LeuCysValAlaLeuAlaIleIleGlyMetMetTrpLeuArgPheLysLysProGluLeu	430

QY	1480	CATCGTCCTTTCAAGTCGGCCACTGTTCATCCAGCTTTGGTTTTCTTCCATCGCCTCTTC	1533
Db	431	GLUARGPROLLELysValasnLeuAlaLeuProValPheHeIleLeuAlaCysLeuPhe	450
QY	1540	ATGCTTGGCCCTTTCCCTCTATTTCGGACCCATTTAGTACAGGATTTGGCTTCGTATCATCACT	1599
Db	451	LeuIleAlaValSerPheThrPlpLysThrProMetGluCysGlyIleGlyPheAlaIleIle	470
QY	1600	CTGACTGGAGTCCTCGTATTATCTCTTTATATATGCGACAAGAACCCAGGTGGTTT	1659
Db	471	LeuSerGlyLeuProValTrpPhePheGlyValTrpTrpLysAsnLysProLysTrpIle	490
QY	1660	AGAATAATGTTCAGAGAAAATAACAGAACATTTACAAATAATACTCGGAAGTTGTACACGAA	1719
Db	491	LeuGlnAlaIlePheSerValThrValLeuCysGlnLysLeuMetGlnValValProGln	510
QY	1720	GAA 1722	
Db	511	Glu 511	
RESULT 6			
Q96QB2		PRELIMINARY;	PRT; 507 AA.
1D	Q96QB2		
AC	Q96QB2; 2001 (TReMBLrel. 19, Created)		
DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)		
DE	Sodium-independent neutral amino acid transporter LAT1.		
GN	LAT1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Nii T., Taketani Y., Tani Y., Ohkido M., Segawa H., Kanai Y., Miyamoto K., Endou H., Takeda E.;		
RT	"Induction of the human sodium-independent neutral amino acid transporter LAT1 (4F2 light chain) in activated T lymphocytes.";		
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB023721; BAB70708.1; JOINED.		
DR	EMBL; AB023713; BAB70708.1; JOINED.		
DR	EMBL; AB023713; BAB70708.1; JOINED.		
DR	EMBL; AB023715; BAB70708.1; JOINED.		
DR	EMBL; AB023716; BAB70708.1; JOINED.		
DR	EMBL; AB023717; BAB70708.1; JOINED.		
DR	EMBL; AB023718; BAB70708.1; JOINED.		
DR	EMBL; AB023719; BAB70708.1; JOINED.		
DR	EMBL; AB023720; BAB70708.1; JOINED.		
DR	InterPro; IPR002293; AA/rel.primase1.		
DR	InterPro; IPR004760; L.AA.transprot.		
DR	InterPro; IPR004841; Permease.		
DR	Pfam; PF00324; aa_permeases; 1.		
DR	TIGRFAMs; TIGR00911; 2A0308; 1.		
SQ	SEQUENCE 507 AA; 55123 MW; DC6C559DAECE21E CRC64;		
Alignment Scores:			
Pred. No.:	4, 81e-96	Length:	507
Score:	1189.50	Matches:	229
Percent Similarity:	65.59%	Conservative:	95
Best Local Similarity:	46.36%	Mismatches:	167
Query Match:	28.87%	Gaps:	3
DB:	4	Indels:	2

[illegible]

```

Db 34 GlyAlaAlaProAlaGlyGlyGly---GluGlyValThrLeuGlnArgAsnIleThrLeu 52
QY 367 CTGAGGGAGTCTCCATTATCATTCATCGGCACCATCATCTGGAGCAGGAATCTTCATCTCCT 426
Db 53 LeuAsnGlyValAlaIleValGlyAlaIleIleGlySerGlyIlePheValThrPro 72
QY 427 AAGGGCGTCTCCAGAACACAGCGCGCATCTGCTGACCATCTGGACGGTGTGT 486
Db 73 ThrGlyValLeuGlyGlyAlaGlySerProGlyLeuAlaLeuValMetTrpAlaAlaCys 92
QY 487 GGGGCTCTGCTACTATTTGGAGCTTTGTCTTTATGCTGAATTTGGAACTATAAAGAAA 546
Db 93 GlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrIleSerLys 112
QY 547 TCTGGAGGTCATTACACATATATTTGGAAGTCTTTGGTCCATTACAGCTTTGTACGA 606
Db 113 SerGlyGlyAspTyrAlaTyrMetLeuAspValTyrGlySerLeuProAlaPheLeuLys 132
QY 607 GTCTGGGTGGAACCTCTCATATATACGCCCTGCAGCTACTGCTGTGATATCCCTGGCATTT 666
Db 133 LeuTrpIleGluLeuLeuValIleArgProSerSerGlnTyrIleValAlaLeuValPhe 152
QY 667 GGACGCTACATCTCGAACCAATTTTATTCATATGTGAATCCCTGAACTTGGCATCAAG 726
Db 153 AlaThrTyrLeuLeuLysProLeuPheProSerCysProValProGluGluAlaAlaLys 172
QY 727 CTCATTACAGCTGTGGGCATAACTAGTGTAGTGTCTCTTAATAGCATGAGTGCACCTGG 786
Db 173 LeuMetAlaCysHisCysValLeuLeuLeuLeuThraValAsnCysTyrSerValLysAla 192
QY 787 AGCGCCGGATCCAGATTTCTTAACTTTTGAAGCTCACAGCAATTTCTGATAATATA 846
Db 193 AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAlaLeuIle 212
QY 847 GTCCCTGGAGTATGACAGTAATTAAGGTCAACGCGAAGCACTTTAAAGCGCCTTTTCA 906
Db 213 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 232
QY 907 -----GGAAGAGATTACAGTATTACGCGGTTCGCCACTGCTTTTATTATGGAATGAT 960
Db 233 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 252
QY 961 GCATATGCTGCGTGTGTTTACCTCACTTTGTTTACTGAGAAGTAGAAAACCTGAAAAA 1020
Db 253 AlaTyrGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
QY 1021 ACCATTCCTCTGCAATATATATCATGCGCATTTGTCACCATTTGGCTATGCTGCACA 1080
Db 273 AsnLeuProLeuAlaIleIleIleSerLeuProIleValThrLeuValTyrValLeuThr 292
QY 1081 AATGTGCGCTTACGACCATTAATGCTGAGGAGCTGCTGCTTTTCAAAATGCAGTGGA 1140
Db 293 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 312
QY 1141 GTGACCTTTCTGAGCGGCTACTGGGAATTTCTCATTAAGCAGTTCCTCATTTGCTGCC 1200
Db 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 332
QY 1201 CTCTCTGCTTTGGCTCCATGACCGGTGTGTGTGTGCTGCTCCAGGTATTCTATGTT 1260
Db 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPheVal 352
QY 1261 GCGTCTCGAGAGGTCACCTCCAGAAATCTCTCCATGATCATGTCGCGAAGCACACT 1320
Db 353 GlySerArgGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372
QY 1321 CCTCTACAGCTGTTATTTGTCACCCCTTTGACAATGATAATGCTCTCTCTCGGAGAC 1380
Db 373 ProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 392
QY 1381 CTCGACAGTCTTTGAAATTTCTCAGTTTGGCCAGGTGGCTTTTATTTGGCTGGCAGTT 1440

```

```

Db 393 IlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIle 412
QY 1441 GCTGGGCTGATTATCTTCGATACAAATGCCAGATATGCTATGCTCTTTCAGGTGCCA 1500
Db 413 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 432
QY 1501 CTGTTTCATCCAGCTTTGTTTTCCTTCACATGCTCTTTCATGTTGCCCTTTCCTCTAT 1560
Db 433 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 452
QY 1561 TCGGACCCATTTCAGTACAGGATTCGGTTCGTCATCACTCTGACTGGAGTCCCTGCTAT 1620
Db 453 LysThrProValGluCysGlyIleGlyPheThrIleIleLeuSerGlyLeuProValTyr 472
QY 1621 TATCTCTTTTATATATGGACAAACCCAGGTGGTTTGAATAATGTGACAGAAATA 1680
Db 473 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuGlnGlyIlePheSerThr 492
QY 1681 ACCAGAACATTACAAATAATACTGGAAGTTGTACCAAGAA 1722
Db 493 ThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 506
RESULT 7
Q92536 PRELIMINARY; PRT; 515 AA.
AC Q92536;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE KIAA0245 protein.
GN KIAA0245
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K., Ohira M., Kawarabayashi Y., Ohara O.,
RA Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT the coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and brain.";
RL DNA Res. 3:321-329(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE=99047611; PubMed=9829974;
RA Torrents D., Estevez R., Pineda M., Fernandez E., Lloberas J.,
RA Shi Y.-B., Zorzano A., Palacin M.;
RT "Identification and characterization of a membrane protein (y+LAT-1)
RT that associates with 4F2hc to encode the amino acid transport activity
RT y+LAT-1. A candidate gene for lysinuric protein intolerance.";
RL J. Biol. Chem. 273:32437-32445(1998).
RL EMBL; D87432; BAAL3376.1;
DR InterPro; IPR002293; AA/rel.prmease1.
DR Pfam; PF00324; aa_permeases; 1.
SQ SEQUENCE 515 AA; 56827 MW; 4E92F6E1972351A9 CRC64;
Alignment Scores:
Pred. No.: 3.9e-93 Length: 515
Score: 1156.50 Matches: 218
Percent Similarity: 69.16% Conservative: 105
Best Local Similarity: 46.68% Mismatches: 141
Query Match: 28.07% Indels: 3
DB: 4 Gaps: 2
US-09-667-170A-440 (1-2239) x Q92536 (1-515)
QY 334 GAGAAGTCCAGCTCAAGAGGAAGTCACTTTTACTGAGGGAGTCTCCATTATCATTTGGC 393

```

Db 37 GluThrMetGlnLeuLysLysGluIleSerLeuLeuAsnGlyValSerLeuValValGly 56
 Qy 394 ACCATATTGGAGCAGGAATCTTCATCTCTCCTAAGGCGGTGCTCCAGAACACGGCGCAGC 453
 Db 57 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuValHisThrAlaSer 76
 Qy 454 GTGGCATCTCTGACCATCTGGACCGTGTGGGGTCTGTCACATATTGGAGCTTTG 513
 Db 77 TyrGlyMetSerLeuIleValTrpAlaIleGlyGlyLeuPheSerValValGlyAlaLeu 96
 Qy 514 TCTTATGCTGAATTTGGGAACAACATATATAAGAAATCTGGAGGTCAATACACATATATTTG 573
 Db 97 CysTyrAlaGluLeuGlyThrThrIleThrLysSerGlyAlaSerTyrAlaTyrIleLeu 116
 Qy 574 GAAGTCTTGTCATTAACACAGCTTTGTAGGAGCTGGGTGGAGTCTCCATCAATAACGC 633
 Db 117 GluAlaPheGlyGlyPheIleAlaPheIleArgLeuTrpValSerLeuValValGlu 136
 Qy 634 CTGCGAGCTACTGCTGTGATATCCCTGGCATTTGGAGCGTACATCTGGAAACCATTTT 693
 Db 137 ProThrGlyGlnAlaIleAlaIleThrPheAlaAsnTyrIleIleGlnProSerPhe 156
 Qy 694 ATTCATGTAATCCCTGCACTTCCGATCAAGTCAATCAGCTGTCGGGATCACTGTA 753
 Db 157 ProSerCysAspProTyrLeuAlaCysArgLeuLeuAlaAlaCysIleCysLeu 176
 Qy 754 GTGATGCTCTTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 813
 Db 177 LeuThrPheValAsnCysAlaTyrValLysTrpGlyThrArgValGlnAspThrPheThr 196
 Qy 814 TTTTCAAGCTCACAGCAATCTGATATATATAGTCCCTGGAGTATGACGCTAATTA 873
 Db 197 TyrAlaLysValValAlaLeuIleAlaIleValMetGlyLeuValLysLeuCysGln 216
 Qy 874 GGTCAACGCAAGCTTAAAGACCCCTTTTCAGGAGAGATTCAGATTAATACCGGTG 933
 Db 217 GlyHisSerGluHisPheGlnAspAlaPheGluGlySerTrpAspMetGlyAsnLeu 236
 Qy 934 CCACGTGGCTTTTATATGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 993
 Db 237 SerLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTyrAspThrLeuAsnPheVal 256
 Qy 994 ACTGAAGAAGTAGAAACCCCTGAAACCACTTCCCTTGGCAATATGATATATATATG 1053
 Db 257 ThrGluGluIleLysAsnProGluArgAsnLeuProLeuAlaIleGlyIleSerMetPro 276
 Qy 1054 ATGTGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1113
 Db 277 IleValThrLeuIleTyrIleLeuThrAsnValAlaTyrTyrThrValLeuAsnIleSer 296
 Qy 1114 GAGCTGCTGCTTCAATCGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 1173
 Db 297 AspValLeuSerSerAspAlaValAlaValThrPheAlaAspGlnThrPheGlyMetPhe 316
 Qy 1174 TCATTAGCATTCGAGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1233
 Db 317 SerTrpThrIleProIleAlaValAlaLeuSerCysPheGlyGlyLeuAsnAlaSerIle 336
 Qy 1234 TTTTGTCTCTCCAGGTATTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1293
 Db 337 PheAlaSerSerArgLeuPhePheValGlySerArgGluGlyHisLeuProAspLeuLeu 356
 Qy 1294 TCCATGATTATCTCCGCAAGCACACTCTCTACACAGCTGTTTATGTTTGTGTTGTTGTTG 1353
 Db 357 SerMetIleHisIleGluArgPheThrProIleProAlaLeuPheAsnCysThrMet 376
 Qy 1354 ACAATGATAATGCTCTCTGAGACCTGACAGCTCTTTGAAATTCCTCAGTTTGGC 1413
 Db 377 AlaLeuIleTyrLeuIleValGluAspValPheGlnLeuIleAsnTyrPheSerPheSer 396
 Qy 1414 AGTGGCTTTTATGCTGGCTGGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473
 Db 397 TyrTrpPhePheValGlyLeuSerValValGlyGlnLeuTyrLeuArgTrpLysGluPro 416

Qy 1474 GATATGCATCGTCTCTTCAAGGTGCCACTGTTCATCCAGCTTTGTTTCTTCATATGC 1533
 Db 417 LysArgProArgProLeuLysLeuSerValPhePheProIleValPheCysIleCysSer 436
 Qy 1534 CTTCTTCATGTTGCTCTTCCCTCTATTCGGACCATTTAGTACAGGAGTGGCTTCGTC 1593
 Db 437 ValPheLeuValIleValProLeuPheThrAspThrIleAsnSerLeuIleGlyIleGly 456
 Qy 1594 ATCACTCTGACGTGAGTCCCTGCTGCTATATCTC-----TTTATATATATGACCAAGAA 1647
 Db 457 IleAlaLeuSerGlyValProPheTyrPheMetGlyValTyrLeuProGluSerArgArg 476
 Qy 1648 CCGAGTGTGTTAGATAATATGTCAGAGAAATAACAGCA---ACATTAACAATAATATCTG 1704
 Db 477 ProLeuPheIleArgAsnValLeuAlaIleThrArgGlyThrGlnGlnLeuCysPhe 496
 Qy 1705 GAAGTGTGACCAAGAGAGAT 1725
 Db 497 CysValLeuThrGluLeuAsp 503

RESULT 8
 QOWS1 PRELIMINARY; PRT; 510 AA.
 ID QOWS1;
 AC QOWS1; STRAIN=NIH/SWISS; TISSUE=HEART;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Y+LAT1b.
 GN SLC7A7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eumalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIH/SWISS; TISSUE=HEART;
 RX MEDLINE=99094891; PubMed=9878049;
 RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;
 RT "Amino acid transport of y-L-type by heterodimers of 4F2hc/CD98 and
 RT members of the glycoprotein-associated amino acid transporter
 RT family.";
 RL EMBO J. 18:49-57(1999).
 DR EMBL; AJ130943; CAA10255.1; --
 DR MGI; MGI:1337120; SLC7A7.
 DR InterPro: IPR002293; AA/rel_primease1.
 DR InterPro: IPR004841; Permease.
 DR Pfam; PF00324; aa_permeases; 1.
 FT VARIANT 4 4 T -> S.
 SQ SEQUENCE 510 AA; 55691 MW; 9F30FA7BAF126F6C CRC64;

Alignment Scores:
 Pred. No.: 5,06e-91 Length: 510
 Score: 1132.50 Matches: 214
 Percent Similarity: 68.59% Conservative: 107
 Best Local Similarity: 45.73% Mismatches: 142
 Query Match: 27.49% Indels: 5
 DB: 11 Gaps: 3

US-09-667-170A-440 (1-2239) x QOWS1 (1-510)

Qy 334 GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTTACTGAGGGAGTCTCCATTATCATTTGC 393
 Db 30 GluGlnValLysLeuLysLysGluIleSerLeuLeuAsnGlyValCysLeuIleValGly 49
 Qy 394 ACCATCATTTGAGCAGGAATCTTCATCTCTAAGGGCTGCTCCAGACACGGGAGC 453
 Db 50 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuMetTyrSerAlaSer 69
 Qy 454 GTGGCATGCTCTCTGACCATCTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513
 Db 70 PheGlyLeuSerLeuValIleTrpAlaValGlyIlePheSerValPheGlyAlaLeu 89


```

QY 514 TCATTGCTGAATTTGGGAACAACATATATAAGAAATCTGAGGCTCATACACATATATTG 573
Db 514 IleAlaLeuSerGlyLeuPro---PheTyrPhePheIleAlaValProGluHisLys 468
QY 90 CystyrAlaGluLeuGlyThrThrIleLysLysSerGlyAlaSerTyrAlaTyrIleLeu 109
Db 90 CystyrAlaGluLeuGlyThrThrIleLysLysSerGlyAlaSerTyrAlaTyrIleLeu 109
QY 574 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATACGC 633
Db 574 ArgProLeuPheLeuArgArgIleValAlaSerIleThrArgTyrLeuGlnIleLeuCys 488
QY 110 GluAlaPheGlyGlyPheLeuAlaPheIleArgLeuTyrTrpThrSerLeuLeuIleGlu 129
Db 110 GluAlaPheGlyGlyPheLeuAlaPheIleArgLeuTyrTrpThrSerLeuLeuIleGlu 129
QY 634 CCTGACGCTACTGCTGTGATATCCCTGGCATTTGGGACGCTACATCTCTGGAACCATTTT 693
Db 634 CysMetValHisValGluPheThrProValProSerLeuLeuPheAsnGlyValLeu 369
QY 130 ProThrSerGlnAlaValIleAlaIleThrPheAlaAsnTyrMetValGlnProLeuPhe 149
Db 130 ProThrSerGlnAlaValIleAlaIleThrPheAlaAsnTyrMetValGlnProLeuPhe 149
QY 594 ATTCATGTGAATCCCTGCACTTGCATCAAGCTCATACAGCTGTGGGCAATACTGTA 753
Db 594 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuMetTyrSerAlaSer 69
QY 150 ProSerCysGlyAlaProTyrAlaAlaGlyArgLeuLeuAlaAlaCysIleCysLeu 169
Db 150 ProSerCysGlyAlaProTyrAlaAlaGlyArgLeuLeuAlaAlaCysIleCysLeu 169
QY 754 GTGATGCTCTAATAGCAGTGTGACGCTGGAGCGCCGATCCAGATTTCTTAAACC 813
Db 754 ArgProLeuPheLeuArgArgIleValAlaSerIleThrArgTyrLeuGlnIleLeuCys 488
QY 170 LeuThrPheIleAsnCysAlaTyrValLysTrpGlyThrLeuValGlnAspIlePheThr 189
Db 170 LeuThrPheIleAsnCysAlaTyrValLysTrpGlyThrLeuValGlnAspIlePheThr 189
QY 814 TTTTGCAGCTCACAAATTTCTGATAATATATAGTCCCTGGAGTTATGCAGCTAATATA 873
Db 814 TyrAlaLysValLeuAlaLeuIleAlaValIleAlaGlyIleValArgLeuGlyGln 209
QY 874 GGTCAACCGCAGCACTTTAAAGACGCTTTTCAAGAGAGATCAAGTATTACGCGTTG 933
Db 874 GGTCAACCGCAGCACTTTAAAGACGCTTTTCAAGAGAGATCAAGTATTACGCGTTG 933
QY 210 GlyAlaThrAlaAsnPheGluAsnSerPheGluGlySerSerPheAlaMetGlyAspIle 229
Db 210 GlyAlaThrAlaAsnPheGluAsnSerPheGluGlySerSerPheAlaMetGlyAspIle 229
QY 934 CCACTGGCTTTTATTATGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 993
Db 934 CCACTGGCTTTTATTATGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 993
QY 230 AlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTrpAspThrLeuAsnTyrVal 249
Db 230 AlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTrpAspThrLeuAsnTyrVal 249
QY 994 ACTGAAGAAGTGAACACCTGAAACCAATTCCTCCCTTGCATATATATATATATATATAT 1053
Db 994 ThrGluGluIleArgAsnProGluArgAsnLeuProLeuSerIleGlyLeuAsnAlaSerIle 269
QY 1054 ATTTGTCACCATGGCTATGCTGCAGCAATATGTCGCTACTTACGACCATTAATGCTGAG 1113
Db 1054 IleValThrIleLeuTyrLeuThrAsnValAlaValTyrSerValLeuAspIleLys 289
QY 1114 GAGCTGCTCTTCAATGACGTCAGTGCAGTCTTCTTTCGAGCGCTACTGGGAAATTC 1173
Db 1114 GluIleLeuAlaSerAspAlaValAlaValThrPheAlaAspGlnIlePheGlyValPhe 309
QY 1174 TCATTAGCAGTCCGATCTTTGTCCTCTCCCTGCTTGTGCTCCATGAACGTTGTTG 1233
Db 1174 AsnTrpIleProValAlaValAlaPheSerCysPheGlyLeuAsnAlaSerIle 329
QY 1234 TTTGCTGCTCCAGGTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1293
Db 1234 ValAlaAlaSerArgLeuLeuPheValGlySerArgGluGlyHisLeuProAspAlaIle 349
QY 1294 TCCATGATTCATGTCGGCAGCAGTCTCTTACGAGTCTTGTATTTGTCACCCCTTGG 1353
Db 1294 CysMetValHisValGluPheThrProValProSerLeuLeuPheAsnGlyValLeu 369
QY 1354 ACAATGATAATGCTCTCTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413
Db 1354 SerLeuValTyrLeuCysValGluAspIlePheGlnLeuIleAsnTyrTyrSerPheSer 389
QY 1414 AGTGGCTTTTATTTGGCTGGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473
Db 1414 TyrTrpPhePheValGlyLeuSerIleValGlyGlnLeuTyrLeuArgTrpLysAspPro 409
QY 1474 GATATGATTCCTCTTCAAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1533
Db 1474 AspArgProArgProLeuLysLeuSerLeuPhePheProIleIlePheCysLeuCysThr 429
QY 1534 CTCCTCATGTTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1593
Db 1534 IlePheLeuValAlaValProLeuTyrSerAspThrIleAsnSerLeuIleGlyIleGly 449
QY 1594 ATCACTGTGACTGGAGTCCCTGCGCTATATATCTTTTATTATA-----TGGGCAAG 1644

```

```

Db 450 IleAlaLeuSerGlyLeuPro---PheTyrPhePheIleAlaValProGluHisLys 468
QY 1645 AAACCCAGGTTGTTAGATAATGTCAGAGAAATACACAGACATTAACAATAATA--- 1701
Db 469 ArgProLeuPheLeuArgArgIleValAlaSerIleThrArgTyrLeuGlnIleLeuCys 488
QY 1702 CTGGAAGTCTTACCAGAGAAGAT 1725
Db 489 MetSerValAlaAlaGluMetAsp 496

```

RESULT 9

```

Q921K8 PRELIMINARY; PRT; 510 AA.
AC Q921K8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Y-LAT1a (Solute carrier family 7 (Cationic amino acid transporter, y+ system), member 7).
GN SLC7A7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=KIDNEY;
RX MEDLINE=99094891; PubMed=9878049;
RA Pfeiffer R., Rossier G., Spindler B., Meier C., Kuhn L., Verrey F.;
RT "Amino acid transport of y+L-type by heterodimers of 4F2hc/CD98 and
RT members of the glycoprotein-associated amino acid transporter
RT family.";
RL EMBO J. 18:49-57(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AT012754; CAA10170.1; -.
DR EMBL; BC014709; AAI14709.1; -.
DR MGD; MGI:1337120; Slc7a7.
DR InterPro; IPR002293; AA/reel_prmease1.
DR InterPro; IPR004841; Permease.
DR Pfam; PF00324; aa_permeases; 1.
FT VARIANT 4 4
SQ SEQUENCE 510 AA; 55677 MW; 9F30FB1B88126F6C CRC64;

```

Alignment Scores:

```

Pred. No.: 5,06e-91 Length: 510
Score: 1132.50 Matches: 214
Percent Similarity: 68.59% Conservative: 107
Best Local Similarity: 45.73% Mismatches: 142
Query Match: 27.49% Indels: 5
DB: 11 Gaps: 3

US-09-667-170A-440 (1-2239) x Q921K8 (1-510)
QY 334 GAGAAAGTCAGCTGAAGAGAAAGTCACATTACTGAGGGAGTCTCCATATATCATTTGCG 393
Db 30 GluGlnValLysLeuLysLysGluIleSerLeuLeuAsnGlyValCysLeuIleValGly 49
QY 394 ACCATCATTTGAGCAGGAATCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 453
Db 50 AsnMetIleGlySerGlyIlePheValSerProLysGlyValLeuMetTyrSerAlaSer 69
QY 454 GTGGCATGCTCTGACCACTGACCGGTGTGGGGTGTGGGGTGTGGGGTGTGGGGTGTGGGG 513
Db 70 PheGlyLeuSerLeuValIleTrpAlaValGlyIlePheSerValPheGlyAlaLeu 89
QY 514 TCATTGCTGAATTTGGGAACAACATATATAAGAAATCTGAGGCTCATACACATATATTG 573
Db 90 CystyrAlaGluLeuGlyThrThrIleLysLysSerGlyAlaSerTyrAlaTyrIleLeu 109

```

QY 574 GAAGTCTTTGGTCCATTACAGCTTTTGTACAGCTGGTGGAACTCTCTCATATACGC 633
 DB 110 GluAlaPheGlyGlyPheLeuAlaPheLeuArgLeuTrpThrSerLeuLeuLeuLeuGlu 129
 QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGAGCTGTACATTCCTGGAACCATTTT 693
 DB 130 ProThrSerGlnAlaValAlaIleThrPheAlaAsnTyrMetValGlnProLeuPhe 149
 QY 694 ATTCATAGTGAATCCCTCAAGCTGGCATCAAGCTCATATACAGCTGTGGGCATATACTGA 753
 DB 150 ProSerCysGlyAlaProTyrAlaAlaGlyArgLeuLeuAlaAlaIleCysLeu 169
 QY 754 GTGATGCTCTAATACATGAGTGTGAGTGGAGCCCGGATCCAGATTTTCTTAACC 813
 DB 170 LeuThrPheIleAsnCysAlaTyrValTyrTrpGlyThrLeuValGlnAspIlePheThr 189
 QY 814 TTTTGGCAAGCTCACAGCAATCTGATATATATAGTCCCTGGAGTGTATGACAGCTAATAA 873
 DB 190 TyrAlaIleValLeuAlaLeuIleAlaValIleIleAlaGlyIleValArgLeuGlyGln 209
 QY 874 GGTCAACGCGAAGCTTTAAAGAGCGCTTTTCAGGAAGAGATTCAAGTATTACCGGGTTG 933
 DB 210 GlyAlaThrAlaAsnPheGluAsnSerPheGluGlySerPheAlaMetGlyAspIle 229
 QY 934 CCAGTGGCTTTTATTTATGGAATGATGATGATGCTGCTGGTGTATACCTCAACTTTGTT 993
 DB 230 AlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSerGlyTrpAspThrLeuAsnTyrVal 249
 QY 994 ACTGAAGAGTAGAAGCAACCTTGAAAAACCATTCCTCCCTGCAATATATATATCATGGCC 1053
 DB 250 ThrGluGluIleArgAsnProGluArgAsnLeuProLeuSerIleGlyIleSerMetPro 269
 QY 1054 ATTTGACCATTTGGCTATGCTGACAAATGTCCTACTTTAGCACCATTAATGCTGAG 1113
 DB 270 IleValThrIleIleTyrLeuLeuThrAsnValAlaTyrTyrSerValLeuAspIleLeu 289
 QY 1114 GAGTGTGCTTCAAAATGAGTGCAGTGCAGTCTTTCTGAGCGGTACTGGGAAATTC 1173
 DB 290 GluIleLeuAlaSerAspAlaValAlaValThrPheAlaAspGlnIlePheGlyValPhe 309
 QY 1174 TCATTAGCAGTTCGATCTTTTGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1233
 DB 310 AsnTyrPheIleProValAlaValAlaPheSerCysPheGlyGlyLeuAsnAlaSerIle 329
 QY 1234 TTTGCTGCTCCAGGTATTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1293
 DB 330 ValAlaAlaSerArgLeuLeuPheValGlySerArgGlyGlyHisLeuProAspAlaIle 349
 QY 1294 TCATGATTCATGTCGCGAAGCACACTCTCTACAGCTGTTATTTGCTGACCCCTTTG 1353
 DB 350 CysMetValHisValGluArgPheThrProValProSerLeuLeuPheAsnGlyValLeu 369
 QY 1354 ACAATGATATGCTCTCTCTGAGACCTCGACAGCTTTTGAATTTCTCAGTTTTCCT 1413
 DB 370 SerLeuValTyrLeuCysValGluAspIlePheGlnLeuIleAsnTyrTyrSerPheSer 389
 QY 1414 AGTGGCTTTTATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
 DB 390 TyrTrpPhePheValGlyLeuSerIleValGlyGlnLeuTyrLeuArgTrpIleAspPro 409
 QY 1474 GATATGATGCTGCTTCAAGTGGCAGCTGTTTCATCCAGCTGTTGTTTCTTCCATATG 1533
 DB 410 AspArgProArgProLeuLysLeuSerLeuPhePheProIleIlePheCysLeuCysThr 429
 QY 1534 CTCCTTCATGCTGCTTCCCTCTATTCGGACCATTTAGTACAGGATGGCTTCGCTC 1593
 DB 430 IlePheLeuValAlaValProLeuTyrSerAspThrIleAsnSerLeuIleGlyIleGly 449
 QY 1594 ATCAGTCTGATGAGTCCCTGGTATTTATCTCTTTATTTATA- - - - -TGGGACAA 1644
 DB 450 IleAlaLeuSerGlyLeuPro- - -PheTyrPhePheIleIleArgValProGluHisLys 468
 QY 1645 AAACCCAGGTGGTTAGATATATGTCAGAGAAATACACAGACATTAACAATAATA- - - 1701

DB 469 ArgProLeuPheLeuArgArgIleValAlaSerIleThrArgTyrLeuGlnIleLeuCys 488
 QY 1702 CTGGAAGTTTACAGAGAAGAT 1725
 DB 489 MetSerValAlaAlaGluMetAsp 496
 RESULT 10
 Q9R0S5 PRELIMINARY; PRT; 512 AA.
 ID Q9R0S5;
 AC Q9R0S5;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Y+L amino acid transporter 1.
 GN RY+LAT1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN SEQUENCE FROM N.A.
 RP TISSUE=KIDNEY;
 RA Kanai Y., Fukasawa Y., Segawa H., Endou H.;
 RT "Characterization of a system y+L amino acid transporter.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB020520; BAA87325.1;
 DR InterPro: IPR002293; AA/rel_permease1.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 SQ SEQUENCE 512 AA; 55641 MW; 708905B2ACB0A824 CRC64;
 Alignment Scores:
 Pred. No.: 6, 2e-91 Length: 512
 Score: 1131.50 Matches: 217
 Percent Similarity: 67.37% Conservative: 103
 Best Local Similarity: 45.68% Mismatches: 152
 Query Match: 27.46% Indels: 3
 DB: 11 Gaps: 2
 US-09-667-170A-440 (1-2239) x Q9R0S5 (1-512)
 QY 310 GGAAGACAGCCCTTTCAGAGAGAGAAAGTCCAGCTCAAGAGAAAGTCACTTACGTG 369
 DB 24 GlyAspGlyAlaGlyProAlaGluGlnValLysGlyLysGlyLysLeuLeu 43
 QY 370 AGGGAGTCTCCATTATTCATTTGGCACCATTCATGGCAGGAGTCTTCTCTCTCTCTCT 429
 DB 44 AsnGlyValCysLeuIleValGlyAsnMetIleGlySerGlyIlePheValSerProLys 63
 QY 430 GGCCTGCTCCAGAACACGGCAGCGTGGGATCTCTCTGACCATCTGGAGCGGTGTGGG 489
 DB 64 GlyValLeuMetTyrSerAlaSerPheGlyLeuSerLeuValIleTrpAlaValGlyGly 83
 QY 490 GTCTGTCACTATTGGAGCTTGTCTTATGCTGAATTTGGGAACAACTATAAGAAATCT 549
 DB 84 IlePheSerValPheGlyAlaLeuCysTyrTrpAlaGluLeuGlyThrThrIleLysLysSer 103
 QY 550 GGAGGTCAATACATATATTTTGGAGTCTTTGGTCCATTTACCACTTTGTACAGTC 609
 DB 104 GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu 123
 QY 610 TGGGTGAACTCTCATATACCCCTGACAGTCTGCTGATATCCCTGGCATTTGGA 669
 DB 124 TrpThrSerLeuLeuIleIleGluProThrSerGlnAlaValIleAlaIleThrPheAla 143
 QY 670 CGCTACATCTGGAACCATTTTATTCAATGTAATCCCTGAACCTGCGATCAAGCTC 729
 DB 144 AsnTyrMetValGlnProLeuPheProSerCysGlyAlaProTyrAlaAlaGlyArgLeu 163
 QY 730 ATTACAGCTGTGGCATACTAGTGTCTTAAATAGCATGTGTGCTGCTGCTGCTGCTGCTG 769
 DB 164 LeuAlaAlaLacysIleCysLeuLeuThrPheIleAsnCysAlaTyrValIleStrpGly 183

QY 790 GCCGGATCCAGATTTCTTAACTTTTGAAGCTCAGCAATCTGTAATATATAGTC 849
Db 184 ThrLeuValGlnAspIlePheThrTyrAlaLysValLeuAlaLeuIleAlaValIle 203
QY 850 CTGGAGTTATGACGATTAATAAGGTCAACAGCAAGACTTAAAGACGCTTTTCAGGA 909
Db 204 AlaGlyIleValArgLeuGlyGlnGlyAlaThrThrAsnPheGluAspSerPheGluGly 223
QY 910 ACAGATTCACGATTTACCGGTTGCCACTGCTCTTTTATTATGAAATGATGATGCT 969
Db 224 SerSerPheAlaMetGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer 243
QY 970 GCGTGGTTTACCTCAACTTTGTCTACTGAGAAAGTACAGAACTGAGAAACCAATCCC 1029
Db 244 GlyTyrPaspThrLeuAsnTyrValThrGluGluIleArgAsnProGluArgAsnLeuPro 263
QY 1030 CTGTCAATATCATCATCCAGCATTCATGTCACCAATGGCTATGCTGTCGAAATGGCC 1089
Db 264 LeuSerIleGlyIleSerMetProIleValThrIleIleTyrLeuLeuThrAsnValAla 283
QY 1090 TACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAATGCAAGTGGCAGTGCCTTT 1149
Db 284 TyrTyrSerValLeuAspIleLysAspIleLeuAlaSerAspAlaValAlaValThrPhe 303
QY 1150 TCTGAGCGCTACTGGGAATTTCTCATAGCAGTTCCTGCTGCTGCTGCTGCTGCTG 1209
Db 304 AlaAspGlnIlePheGlyIlePheAsnTyrThrIleProLeuAlaValAlaLeuSerCys 323
QY 1210 TTTGGCTCCATACGCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1269
Db 324 PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuLeuValGlySerArg 343
QY 1270 GAGGTCACCTCCAGAAATCTCTCATGATTCATGTCGCAAGCACACTGCTCTACCA 1329
Db 344 GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro 363
QY 1330 GCTGTTATGTTTGGACCTTTGACAAATGATATGCTCTCTCTGAGACCTCGACAGT 1389
Db 364 SerLeuLeuPheAsnGlyIleLeuAlaLeuValTyrLeuCysValGluAspIlePheGln 383
QY 1390 CTTTGTGAATTCCTCAGTTTGGCAGGTGCTTTTATTGGCTGGCAGTTCCTGGGCTG 1449
Db 384 LeuIleAsnTyrTyrSerPheSerTyrTyrPhePheValGlyLeuSerIleValGlyGln 403
QY 1450 ATTATCTTCGATACAAATCCAGATATGCAATGCTGCTCTCAAGTGCCACTGCTCATC 1509
Db 404 LeuTyrLeuArgTyrPlysGluProAspArgProArgProLeuLysLeuSerLeuPhePhe 423
QY 1510 CCAGCTTTGTTTCTTCATGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1569
Db 424 ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr 443
QY 1570 TTTAGTACAGGATTCGCTGCTATCATCTGCTGAGTGCCTGCTGCTGCTGCTGCTGCT 1629
Db 444 IleAsnSerLeuIleGlyIleGlyIleAlaLeuSerGlyLeuProPheTyrPheLeuIle 463
QY 1630 ATT-----ATATGGACAAAGAACCCAGGTGTTTACAAATATGCTCAGAGAAATACC 1683
Db 464 IleArgValProGluHisIysArgProLeuCysLeuArgArgIleValAlaSerThrThr 483
QY 1684 AGAACATTACAAATAATA---CTGGAAGTTGTACCAGAAAGAGAT 1725
Db 484 ArgTyrLeuGlnIleIleCysMetSerValAlaAlaGluMetAsp 498
RESULT 11
Q90266 PRELIMINARY; PRT; 512 AA.
AC Q90266;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2002 (TREMBlrel. 20, Last annotation update)
DE Amino acid transporter yLATL.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RA Ng A.M.L., Yao S.Y.M., Cheeseman C.I., Young J.D.;
RT "cDNA encoding rat jejunal amino acid transporter y-LATL";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF200684; AAR07216.1; -;
DR InterPro: IPR002293; AA/rei_primeasel.
DR InterPro: IPR004841; Permease.
DR Pfam: PF00324; aa_permeases; 1.
SQ SEQUENCE 512 AA; 55684 MW; 708905B2ACB2130B CRC64;

Alignment Scores:
Pred. No.: 6,2e-91 Length: 512
Score: 1131.50 Matches: 217
Percent Similarity: 67.37% Conservative: 103
Best Local Similarity: 45.68% Mismatches: 152
Query Match: 27.46% Indels: 3
Db: 11 Gaps: 2

US-09-667-170A-440 (1-2239) x Q90266 (1-512)
QY 310 GGAAGAGAGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAGAAAGTCTACTG 369
Db 24 GlyAspGlyAlaGlyProAlaAlaGluGlnValLysLeuLysLysGluIleSerLeuLeu 43
QY 370 AGGGAGCTCTCCATTATCATTTGCGACCATCATTTGGAGCAGGAATCTTCATCTCTCT 429
Db 44 AsnGlyValCysLeuIleValGlyAsnMetIleGlySerGlyIlePheValSerProlys 63
QY 430 GCGCTGCTCCAGAACACGCGGCGGCTGCTCTGACCATCTGACCATCTGGAGCGTGTGG 489
Db 64 GlyValLeuMetTyrSerAlaSerPheGlyLeuSerLeuValIleThrAlaValGlyGly 83
QY 490 GTCCTGCTACATTTTGGAGCTTTGCTTATGCTGAATGGGAACACTATAAGAAATCT 549
Db 84 IlePheSerTyrAlaPheGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleLysLysSer 103
QY 550 GGAGGTCATTACACATATATTTTGGAGTCTTTGGGCCATTACCAGCTTTTCTACGAGTC 609
Db 104 GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu 123
QY 610 TGGGTGGAAGTCTCATATAGCCCTGCGAGTCTGCTGATATCCCTGCAATTTGGA 669
Db 124 TrpThrSerLeuLeuIleIleGluProThrSerGlnAlaValIleAlaIleThrPheAla 143
QY 670 CGCTACATTTCTGGAACCATTTTATTCATCTGAATCCCTGAACTTGGCATCAAGCTC 729
Db 144 AsnTyrMetValGlnProLeuPheProSerCysGlyAlaProTyrAlaAlaGlyArgLeu 163
QY 730 ATTACAGCTGGGCGATCAACTGCTAGTGTGCTTAAATAGCATGAGTGTGCTGAGTGG 789
Db 164 LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyrValLysTrpGly 183
QY 790 GCGCGGATCCAGATTTTCTTAACTTTTTCAGAGCTCAGCAATTTCTGATTAATATGCT 849
Db 184 ThrLeuValGlnAspIlePheThrTyrAlaLysValLeuAlaLeuIleAlaValIle 203
QY 850 CCTGGAGTTATGCAAGTAAATAAAGCTCAAAAGCGAGAACTTTAAAGAGCGCTTTTCAGGA 909
Db 204 AlaGlyIleValArgLeuGlyGlnGlyAlaThrThrAsnPheGluAspSerPheGluGly 223
QY 910 AGAGATTCAAGTATTACGGGTTGCCACTGCTGCTTTTATTATGGAATGCTATGCTATGCT 969
Db 224 SerSerPheAlaMetGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer 243
QY 970 GCGTGGTTTACCTCAACTTTGCTTACTGAGAAAGTACAGAACTGAGAAACCAATCCC 1029
Db 244 GlyTyrPaspThrLeuAsnTyrValThrGluGluIleArgAsnProGluArgAsnLeuPro 263

Db 108 MetIleThrLysSerGlyAlaAspTyrAlaTyrIleMetGluThrPheGlyProPheMet 127
 QY 595 GCTTTGTACAGCTGGTGGGAATCTCATATACGCGCTCGACTACTGCTGTGATA 654
 Db 128 AlaPheIleArgLeuThrIleGlyCysMetIleValArgProCysSerGlnAlaIleVal 147
 QY 655 TCCCTGGCATTTGGACGCTACATCTCTGGAACCACTTTTATTCATGTAATCCCTGAA 714
 Db 148 AlaLeuThrPheSerThrTyrValLeuLysProPhePheProGluCysThrProProGlu 167
 QY 715 CTGCGATCAAGCTCATACAGCTGGGCGATCAACTGTAGTGTCTCTAAATAGCATG 774
 Db 168 AspSerAlaArgLeuLeuAlaValCysCysIleLeuValLeuThrLeuIleAsnCysTrp 187
 QY 775 AGTCTAGCTGGAGCGCGCGATTCACAGATTTCTTAACCTTTTGGCAAGCTCACACAAT 834
 Db 188 AspValLysTrpAlaThrAlaValGlnAspIlePheThrTyrAlaLysLeuLeuAlaLeu 207
 QY 835 CTGATATTATAGTCCCTGGAGTATGACGCTAAATTAAGGTCAACAGCAGAACCTTAA 894
 Db 208 PheIleIleAlaThrGlyValTyrGlnLeuTyrLeuGlyAsnThrGlnTyrPhe--- 226
 QY 895 GAGCGCTTTTCAGGAAGAGATTCAGATATACGCGTTCCCACTGGCTTTTATATGGA 954
 Db 227 ---ThrPheGluAsnThrAspThrLysValThrSerIleAlaLeuSerPheTyrSerGly 245
 QY 955 ATGTATGCATATGCTGGCTTACCTCAACTTTGTTACTGGAAGAGTACAGAAACCT 1014
 Db 246 LeuPheAlaTyrAsnGlyTrpAsnTyrLeuAsnPheIleIleGluGluLeuLysAspPro 265
 QY 1015 GAAAAACCAATTCCTTCATATGTATATCATTCATGTCACCATTTGGCTATGTG 1074
 Db 266 ValLysAsnLeuProArgAlaIleAlaIleSerCysThrLeuValThrIleValTyrVal 285
 QY 1075 CTGACAAATGTGCTTACTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAATGCA 1134
 Db 286 MetAlaAsnValSerPheTyrThrIleLeuSerProAspGluValMetGlySerSerAla 305
 QY 1135 GTGGCAGTGCCTTTCTGAGCGGTACTGGGAATTTCTCATAGCAGTTCCGACTTT 1194
 Db 306 ValAlaValThrTyrAlaGluArgAlaPheGlyMetLeuAlaThrIleProValPhe 325
 QY 1195 GTTGCCCTCTCTGCTTGGCTCCATGACGCTGGTGTGCTGCTCCAGTTATTC 1254
 Db 326 ValAlaLeuSerThrPheGlyAlaValAsnGlyIleLeuLeuThrSerSerArgLeuPhe 345
 QY 1255 TATGTTGCTCTCGAGGGTCACTTCCAGAAATCTCTCCATGATCATGTCGCCGAG 1314
 Db 346 TyrAlaGlyAlaAsnAsnGlyGlnMetProGluIleLeuThrMetIleGlnIleGlnArg 365
 QY 1315 CACACTCTCTACAGCTGTTTATGTTTGGACCCCTTTGACATGATAATGCTCTCTCT 1374
 Db 366 PheThrProThrProAlaValLeuAlaMetAlaLeuLeuSerMetLeuTyrLeuThrVal 385
 QY 1375 GGAGACCTCGACAGCTTTTCAATTTCTCAGTTTCCAGGTTGCTTTATTTGGCTG 1434
 Db 386 SerAspIlePheAlaLeuIleAsnTyrValGlyPheAlaThrIlePheLeuSerIleGlyVal 405
 QY 1435 CGAGTGTGCTGGCTGATTATCTTCGATACAAATGCCAGATGATGCTGCTTTCAAG 1494
 Db 406 AlaValLeuCysLeuProThrLeuArgTrpAlaGlnProAsnLeuProArgProIleArg 425
 QY 1495 GTGGCAGTGTTCATCCAGCTGTTTCTTCTTCATGCTTTCATGCTGCTGCTTCTCC 1554
 Db 426 ValProMetValPheProIleValTyrLeuIleAlaThrIlePheValThrValPro 445
 QY 1555 CTCATTCTGGACCCATTAGTACAGGATTTGGCTTCTGTCATCACTCTGACTGAGTCCCT 1614
 Db 446 MetTyrAlaSerProValGluThrGlyTyrGlyIleLeuMetIleLeuSerSerIlePro 465
 QY 1615 GCGTATTATCTTTATATATGAGCAAGAACCCAGGTGTTTGAATATATGTCAGAG 1674
 Db 466 ValTyrLeuValPheIleAlaThrLysAsnLysProIleThrPheGlnLysThrMetGly 485

QY 1675 AAAATAACCAAGACATTACAAATAATAACTACTGGAAGTTGTA 1713
 Db 486 GlyLeuThrGlnValLeuGlnLysLeuMetMetValVal 498
 RESULT 13
 QY1A7
 ID QY1A7 PRELIMINARY; PRT; 499 AA.
 AC QY1A7;
 DT 01-NOV-1999 (T-EMBLrel. 12, Created)
 DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE MND protein (LD25378p).
 MND OR CG3297.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Martin J.F., Hersperger E., Simcox A., Shearn A.;
 RA "Minidis encodes a component of an endocrine signaling pathway
 RT required for normal imaginal cell proliferation."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,

RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
 Yu C., Lewis S.E., Rubin G.M., Celnik S.,
 Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF03532; AAF49688.1; -;
 DR EMBL; AF139834; AAD39459.1; -;
 DR EMBL; AY069533; AAL39678.1; -;
 DR Flybase; FBgn002778; mnd.
 DR InterPro; IPR002293; AA/rel_primease.
 DR InterPro; IPR004841; Permease.
 DR Pfam; PF00324; aa_permeases; 1.
 SQ SEQUENCE 499 AA; 54350 MW; E2882CC0737CC56P CRC64;

Alignment Scores:
 Pred. No.: 1,09e-88 Length: 499
 Score: 1106.00 Matches: 218
 Percent Similarity: 65.57% Conservative: 102
 Best Local Similarity: 44.67% Mismatches: 156
 Query Match: 26.84% Indels: 12
 DB: 5 Gaps: 3

US-09-667-170A-440 (1-2239) x Q9Y1A7 (1-499)

QY 265 GAGCCACCTGGCAGGAGCGCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGCGCTTT 324
 DB 22 GluProAsnAsnSerThrAlaAspSerGlySerglnGly----- 34
 QY 325 TCAGGAAGAGAGAAAGTCCAGCTGAGGAAGAACTCACTTACTAGGGGAGTCTCCATT 384
 DB 35 SerGly-----ValLeuLeuLysGlnIleGlyLeuLeuAspGlyValAlaIle 51
 QY 385 ATCATTGGCACCATTATGGAGCAGGAATCTTCATCTCTCCTCAAGGCGCTGCCAGAAC 444
 DB 52 IleValGlyValIleValGlySerGlyIlePheValSerProLysGlyValLeuLysPhe 71
 QY 445 ACGGCGAGCGTGGGCTCTCTGTACCACTGAGCGGTGTGGGGTCTGTCTCACTATT 504
 DB 72 SerGlySerglnGlnSerLeuIleValTrpValLeuSerGlyValLeuSerMetVal 91
 QY 505 GGAGCTTTGTCTATGCTGAATGGGAACAATATAAGAAATCTGGAGGCTATACACA 564
 DB 92 GlyAlaLeuLysValAlaGluLeuGlyThrMetIleProLysSerGlyGlyAspTrpAla 111
 QY 565 TATATTTTGGAGTCTTTGGTCCATACAGCTTTTGTACAGCTCTGGTGGGAGTCTCTC 624
 DB 112 TyrIleGlyThrAlaPheGlyProLeuProAlaPheLeuTyrLeuTrpValAlaLeuLeu 131
 QY 625 ATAATAGCCCTGCAGCTACTCTGTGATATCCCTGGCATTTGGAGCGCTACATTCGAA 684
 DB 132 IleLeuValProThrGlyAsnAlaIleThrAlaLeuThrPheAlaIleTyrLeuLys 151
 QY 685 CCATTTTATTTTCAATGCTGAATCCCTGAAGTTCGCACTCAAGCTCATACAGCTGTGGC 744
 DB 152 ProPheTrpProSerCysAspAlaProIleGluAlaValGlnLeuLeuAlaAlaMet 171
 QY 745 ATAATCTAGTGTGCTCAATACATGAGTGTGAGTGTGAGCGCCGATCCAGATT 804
 DB 172 IleCysValLeuThrLeuIleAsnCysTyrAsnValLysTrpValThrArgValThrAsp 191
 QY 805 TTCTTTAATCTTTGCAAGCTCAGCAATTCGATAATATATAGTCCCTGGAGTTATGAG 864
 DB 192 IlePheThrGlyThrLysValValAlaLeuLeuValIleValGlyAlaGlyValTrpTrp 211
 QY 865 CTAATTAAGGTCAAGCGAGAACTTTAAGAGCGCTTTTCAGGAAGAGATTCAGATT 924
 DB 212 LeuPheAspGlyAsnThrGluHisTrpAspAsnProPheSerGlyGlyLeuGlnAspPro 231
 QY 925 ACGGGTGTGCACGCGCTTTTATATGGAATGATGATGATGATGATGATGATGATGATGAT 984
 DB 232 GlyTyrIleAlaLeuAlaPheTyrSerGlyPhePheSerTyrSerGlyTrpAsnTyrLeu 251
 QY 985 AACTTTGTACTGAAGAGTAGAAACCTGAAACCAACCTCCCTTCGATATGATATGATAT 1044
 DB 252 AsnPheValThrGluGluLeuLysAspProTyrArgAsnLeuProLysAlaIleCysIle 271

QY 1045 TCCATGGCCATTCTCACCATTGGCTATGTGTGACAAATGTGGCTACTTTACGACCATT 1104
 DB 272 SerMetProValValThrValIleTyrMetIleThrAsnIleAlaTyrPheSerValLeu 291
 QY 1105 AATGTGTGAGGAGCTGCTGCTTTCAATGAGTGGCAGTACGCTTTTCTGAGGGGTACTG 1164
 DB 292 SerProAspGluIleLeuSerSerAspAlaValAlaValThrPheGlyAspLysMetLeu 311
 QY 1165 GGAATTTTCAATAGCAGTTCGATCTTTGTGGCTCTCCTGCTTGGTGGCTCCATGAAC 1224
 DB 312 GlyTyrMetSerTrpIleMetProPheAlaValAlaCysSerThrPheGlySerLeuAsn 331
 QY 1225 GGTGTGTGTGTCTGCTCCAGGTATTCTATGTGTGCTCTCGAGAGGTCACCTTCCA 1284
 DB 332 GlyAlaIlePheAlaSerArgLeuPhePheValGlyAlaArgAsnGlyHisLeuPro 351
 QY 1285 GAAATCCTCTCCATGATCATGTCCGAGCAGCAGTCTCTACAGCTCTTATTTGTTG 1344
 DB 352 AlaAlaIleSerLeuIleAsnValCysLeuThrProValProSerLeuIlePheLeu 371
 QY 1345 CACCTTTGACATGATAATGCTCTCTCTGGAGACTCGACAGTCTTTTGAATTCCTC 1404
 DB 372 GlyValLeuThrIleLeuLeuLeuPheIleGluAspValTyrValLeuIleAsnTyrVal 391
 QY 1405 AGTTTTCAGGTCGCTTTTATTTGGCTGGCAGTGTCTGGCTGATTTATCTTCGATAC 1464
 DB 392 SerTyrValGluAlaLeuPheThrLeuIleSerValSerGlyLeuLeuTrpMetArgTyr 411
 QY 1465 AATGCCCATATGATGCTGCTTTCAAGTGCCAGTGTTCATCCAGCTTTGTTTTC 1524
 DB 412 LysGlnProLysThrGluArgProIleLysValAsnLeuAlaLeuProIleIleTyrLeu 431
 QY 1525 TTCACATGCTCTTCATGTTGGTGGCTTCCCTCTATTCGAGCCATTTAGTACAGGATT 1584
 DB 432 IleValCysLeuPheLeuValIlePheSerCysThrGlnThrProTyrValValGlyIle 451
 QY 1585 GCTTCGTCATCAGCTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
 DB 452 GlyThrIleIleLeuSerGlyIleProValTyrIleLeuThrIle-----HisLys 469
 QY 1645 AAACCCAGGTGTTAGAAATATGTCAGAGAAATATACAGAACATTCACAAATTAATCTG 1704
 DB 470 ProValLysTrpLeuAlaAspThrSerGlnAlaIleAsnLeuTrpCysSerLysPhePhe 489
 QY 1705 GAAGTGTACCAAGAGAGATAAG 1728
 DB 490 IleCysMetProAsnGlnGluLys 497

RESULT 14
 Q9N1Q4 PRELIMINARY; PRT; 535 AA.
 ID Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4;
 AC Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4; Q9N1Q4;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE 4F21c-5.
 DE 4F21c-5.
 GN 4F21c-5.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OC NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20473842; PubMed=10631289;
 RA Rajan D.P., Kekuda R., Huang W., Devoe L.D., Leibach F.H.,
 RA Prasad P.D., Ganapathy V.;
 RT "Cloning and functional characterization of a Na(+)-independent,
 RT broad-specific neutral amino acid transporter from mammalian
 RT intestine";
 RL Biochim. Biophys. Acta 1463:6-14(2000).
 DR EMBL: AF170106; AAF26282.1; -;
 DR InterPro: IPR002293; AA/rel_primease1.

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulo G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., M.G.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195(2000).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Parasas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003759; AAF56667.1;
DR EMBL; AY060663; AAL28211.1;
DR FlyBase: FBgn0039487; CG6070.
DR InterPro: IPR002293; AA/re1.prmease1.
DR InterPro: IPR004841; Permease.
DR Pfam: PF00324; aa_permeases; 1.
SQ SEQUENCE 517 AA; 56522 MW; F1B80A7F51F18C8 CRC64;

Alignment Scores:

Pred. No.: 2,3e-87 Length: 517
Score: 1091.00 Matches: 210
Percent Similarity: 64.86% Conservative: 113
Best local Similarity: 42.17% Mismatches: 167
Query Match: 26.48% Indels: 8
DB: 5 Gaps: 4

US-09-667-170a-440 (1-2239) x Q9VB75 (1-517)

QY 247 CCTTCCTGGGCAACAGAGCCACT-----GGCAGGAGCCCTTTTCAGGAAGA 297
Db 19 ProSerSerGlyAsn-----ProProSerSerIleAlaGlnProAlaAsnProAla 36
QY 298 GAGCGCTTTCAGGAGAGAC-----GCCCTTTTCAGGAGAGAGAAAGTCAGCTGAAG 351
Db 37 GluLysAlaGlnCysArgGluGlySerAlaGluSerAspSerSerArgValValIleLys 56
QY 352 AGGAAAGTCACCTTTACTGAGGGAGTCCCAATTATCATTTGGCACCACCATTTGGAGCAGGA 411
Db 57 LysGlnLeuGlyLeuLeuGluGlyValAlaIleIleLeuGlyIleIlePheGlySerGly 76
QY 412 ATCTTCATCTCTTAAGGGCTGTCTCCAGACACGGGAGCGGTGGGATGCTCTGACC 471
Db 77 IlePheValSerProLysGlyValIleArgGluValGluSerValGlyAlaSerLeuVal 96
QY 472 ATCTGCAGCGTGTGGGCTCTGTCACATTTTGGAGCTTCTCTATGCTGAAATGGGA 531
Db 97 IletrpValLeuCysGlyLeuLeuSerMetIleGlyAlaLeuCysTyrAlaGluLeuGly 116
QY 532 ACAACTATAAAGAAATCTGGAGGTCTATTACACATATATTTTGGAAAGTCTTTGGTCCATA 591
Db 117 ThrAlaIleProLysSerGlyGlyAspTyrAlaTyrIlePheGluAlaTyrGlySerLeu 136

QY 592 CCAGCTTTTGTACGAGTCTGGGTGGAACTCCTCATAATACGCCCTGACGCTACTGCTGTG 651
Db 137 ProAlaPheLeuTyrLeuTyrAspAlaMetIlePheValProThrThrAsnAlaIle 156
QY 652 ATATCCCTGGCATTTGGAGCGTACATTTCTGGAACCACTTTT---ATTCAATGTGAATC 708
Db 157 MetGlyLeuThrPheAlaSerTyrValLeuGluProPhePheGlyGlyAlaCysGluIle 176
QY 709 CCGTGAATTCGATCAAGCTCATACAGCTGTGGCATTAACCTAGTGTAGTGTCTTAAT 768
Db 177 ProLysIleAlaLeuGlnLeuLeuAlaIleThrIleCysPheLeuThrTyrLeuAsn 196
QY 769 AGCATGAGTGTACGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACA 828
Db 197 SerTyrTyrMetLysValThrThrLysMetGlnAsnValIleMetPheThrIleAla 216
QY 829 GCAATTCCTAATAATATAGTCCCTGGAGTTATGACAGTAAATTAAGGTCAACGAGAAC 888
Db 217 AlaLeuValLeuIleIleLeuValGlyLeuValTyrMetMetMetGlyAsnValGluAsn 236
QY 889 TTTAAAGACGCTTTTCAGGAAGAGATTCACAGTATTACGCGTGTGCCACTGCTTTTAT 948
Db 237 PheThrArgProPheAspAsnThrGluThrAspProGlyLysMetSerValAlaPheTyr 256
QY 949 TATGGAATCATCATATGCTGCTGGTGTTTTACCTCAACTTTTGTACTGAAGAAGTAGAA 1008
Db 257 SerGlyIlePheSerTyrAlaGlyTrpAsnTyrLeuAsnPheMetThrGluGluLeuArg 276
QY 1009 AACCTTGAAGAAACCAATCCCTTCCATATGATATATATCATGATGATGATGATGATGAT 1068
Db 277 AspProTyrArgAsnLeuProArgAlaIleTyrIleSerLeuProLeuValThrGlyIle 296
QY 1069 TATGCTGCTGACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1128
Db 297 TyrValLeuAlaAsnValAlaTyrLeuAlaValLeuSerProSerGluMetIleAlaSer 316
QY 1129 AATGAGTGGCAGTGACCTTTCTGAGCGCTACTGAGGAAATTTCTCATAGTAGTTCG 1188
Db 317 AsnAlaIleAlaValThrPheGlyAspLysIleLeuGlyValPheSerLeuIleIlePro 336
QY 1189 ATCTTTGTTGCCCTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
Db 337 LeuMetValAlaIleSerAlaPheGlyGlyLeuSerValHisIleMetThrSerSerArg 356
QY 1249 TTATTTCTATCTTCGCTGCTGAGAGGCTCACCTTCCAGAAATCTCTCCATGATTCATGTC 1308
Db 357 IleCysPheValGlyAlaArgAsnGlyHisMetProAlaIleLeuSerHisIleSerVal 376
QY 1309 CGCAAGCACACTCCTCTACAGCTGTATTGTTTGCACCCCTTTGACAAATGATGATGCTC 1368
Db 377 LysSerTyrThrProLeuProSerLeuValPheLeuCysPheLeuSerIleValMetLeu 396
QY 1369 TTCTCTGGAGACTCGACACTCTTTTCAATTTCTCAGTTTCCAGTTTCCAGTTTCCAGTTT 1428
Db 397 ValValSerAspValTyrValLeuIleThrTyrAlaSerIleValGluSerPhePheIle 416
QY 1429 GGGCTGCCAGTCTGCTGGGCTGATTTATCTTCGATACAAATGCCCAGATATGATCCTCT 1488
Db 417 MetLeuSerValSerAlaValLeuTyrPheArgTyrThrArgProCysMetGluArgPro 436
QY 1489 TTCAGGTGCGCACTGTTTCATCCAGCTTTTCTTCTTCCATCCATCCCTCTTCTGTTGGC 1548
Db 437 IleLysValAlaMetTrpIleProAlaLeuPheValIleValCysAlaPheLeuValVal 456
QY 1549 CTTTCCCTCTATTCGGGACCCCATTTAGTACAGGATTTGGCTTCGTCATCATCTCTGCTGGA 1608
Db 457 ValProIleTyrValAlaProTyrGluValGlyMetGlyValLeuIleThrIleGly 476
QY 1609 GTCCCTCGGTATATCTCTTTATATATGGCAAGAAACCCAGGTGGTTAGATAAATG 1668
Db 477 IleProPheTyrTyrValGlyValIleTrpLysAsnLysProLysTrpValGlnSerThr 496
QY 1669 TCAGAGAAATAACAGAACATTACAAATAATCTGGAAGTTGTACCAAGAA 1722

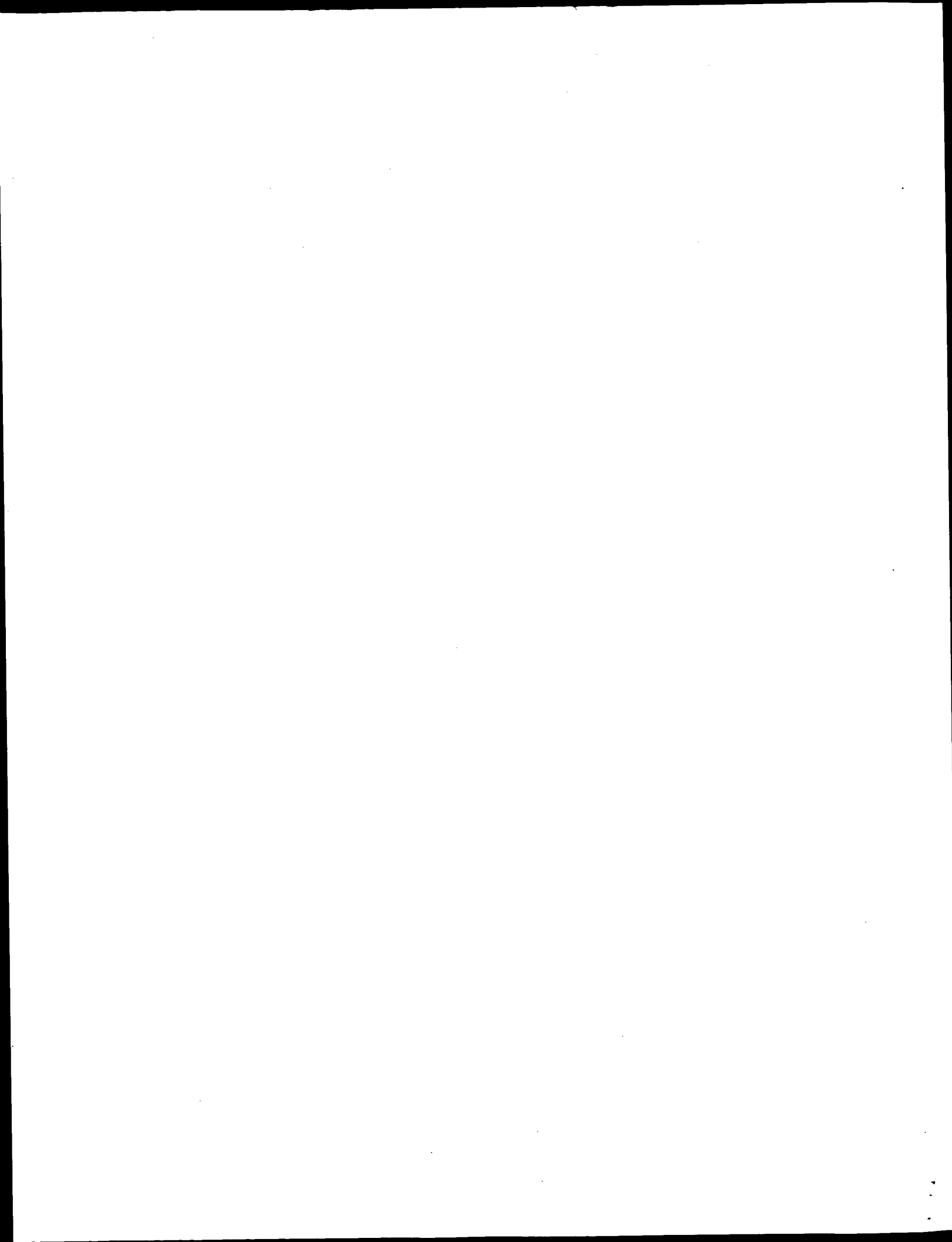
Thu Apr 17 09:48:27 2003

us-09-667-170a-440.n2p.rspt

Page 17

Db 497 IleAspSerValThrPheThrCysGlnLysLeuPheMetSerAlaLysGlu 514
::: ::::: ||| ||| :::
|||||

Search completed: April 16, 2003, 16:14:05
Job time : 113.5 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 10:55:43 ; Search time 93 Seconds
(without alignments)
7383.325 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0
Searched: 441362 seqs, 153338381 residues

Word size : 0
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	27	1.2	737	2	US-08-861-549-2
2	24	1.1	40000	4	US-09-780-049-18
3	24	1.1	162450	4	US-09-345-882-1
4	23	1.0	72928	3	US-09-009-913-1
5	22	1.0	333	4	US-09-018-584A-27
6	22	1.0	350	2	US-08-623-908A-17
7	22	1.0	11288	4	US-08-646-301A-1
8	22	1.0	11288	4	US-08-481-968A-4
9	22	1.0	11288	4	US-08-154-712B-4
10	22	1.0	36651	4	US-09-738-894A-3
11	22	1.0	246240	2	US-08-724-394A-20
12	22	1.0	246240	2	US-08-724-394A-21
13	22	1.0	246240	2	US-08-724-394A-22
14	21	0.9	544	4	US-09-280-116-247
15	21	0.9	579	4	US-09-328-111-687
16	21	0.9	668	4	US-09-328-111-687
17	21	0.9	704	4	US-09-475-316A-26
18	21	0.9	950	3	US-08-984-550-1
19	21	0.9	2636	4	US-09-370-807-5
20	21	0.9	2636	4	US-09-921-259-5
21	21	0.9	5510	1	US-08-123-161A-7
22	21	0.9	5510	1	US-08-483-278-7
23	21	0.9	5852	1	US-07-867-106-2
24	21	0.9	8353	3	US-08-611-587-1
25	21	0.9	22067	4	US-09-820-001-3
26	21	0.9	36651	4	US-09-738-894A-3
27	21	0.9	84495	4	US-09-797-906-3

SUMMARIES

28	20	0.9	985	4	US-08-842-306B-1	Sequence 1, Appli
29	20	0.9	985	4	US-08-838-973B-1	Sequence 1, Appli
30	20	0.9	985	4	US-08-771-212A-1	Sequence 1, Appli
31	20	0.9	1229	1	US-08-440-861-1	Sequence 1, Appli
32	20	0.9	1229	1	US-08-433-854-1	Sequence 1, Appli
33	20	0.9	1229	1	US-08-174-745A-1	Sequence 1, Appli
34	20	0.9	1229	2	US-08-195-947-1	Sequence 1, Appli
35	20	0.9	1229	2	US-08-433-885-1	Sequence 1, Appli
36	20	0.9	1229	2	US-08-433-908B-1	Sequence 1, Appli
37	20	0.9	1229	4	US-08-410-614-1	Sequence 1, Appli
38	20	0.9	1242	4	US-08-413-974-1	Sequence 1, Appli
39	20	0.9	1242	4	US-08-434-418-1	Sequence 1, Appli
40	20	0.9	1242	4	US-08-433-288-1	Sequence 1, Appli
41	20	0.9	1242	4	US-08-174-739A-1	Sequence 1, Appli
42	20	0.9	1242	4	US-08-434-256-1	Sequence 1, Appli
43	20	0.9	1304	4	US-09-372-422A-15	Sequence 1, Appli
44	20	0.9	1400	4	US-09-018-584A-35	Sequence 15, Appli
45	20	0.9	1460	4	US-09-257-179-23	Sequence 35, Appli

ALIGNMENTS

RESULT 1
US-08-861-549-2
; Sequence 2, Application US/08861549
; Patent No. 5874246
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Zhang, Hong
; TITLE OF INVENTION: NEW TRANSCRIPTION REPRESSOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/861,549
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0306 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 737 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNORAB01
CLONE: 1620089

US-08-861-549-2
Query Match 1.2%; Score 27; DB 2; Length 737;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2191 GAGGAGAGAAGCTAGAGCTGTTCTTAA 2217
|||||
Db 711 GAGGAGAGAAGCTAGAGCTGTTCTTAA 737

RESULT 2
US-09-780-049-18/c
; Sequence 18, Application US/09780049
; Patent No. 6465250
; GENERAL INFORMATION:
; APPLICANT: J. Brette P. Monia
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PROTEIN PHOSPHATASE 2 CATALYTIC SUBUNIT
; TITLE OF INVENTION: EXPRESSION
; FILE REFERENCE: RTS-0134
; CURRENT APPLICATION NUMBER: US/09/780,049
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 40000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-049-18

Query Match 1.1%; Score 24; DB 4; Length 40000;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 27 CCAGCCTGGGTGACAGTGACATC 50
|||||
Db 15461 CCAGCCTGGGTGACAGTGACATC 15438

RESULT 3
US-09-345-882-1/c
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bougueret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93714
; OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
; FEATURE:
; NAME/KEY: allele

LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817

OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58

FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:
Query Match 1.1%; Score 24; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 27 CCAGCCTGGGTGACAGTGAGACTC 50
|||||
Db 54003 CCAGCCTGGGTGACAGTGAGACTC 53980
RESULT 4
US-09-009-913-1/c
; Sequence 1, Application US/09009913
; Patent No. 6087485
; GENERAL INFORMATION:
; APPLICANT: Axys Pharmaceuticals, Inc.
; TITLE OF INVENTION: Asthma Related Genes
; NUMBER OF SEQUENCES: 339
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Ave, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,913
; FILING DATE: 21-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-4P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3231
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72928 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-009-913-1
Query Match 1.0%; Score 23; DB 3; Length 72928;
Best Local Similarity 100.0%; Pred. No. 1.1;

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11288
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-646-301A-1

Query Match 1.0%; Score 22; DB 4; Length 11288;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AAAAAAGAAAGAAAGAAAGA 92
|||||
DB 834 AAAAAAGAAAGAAAGAAAGA 855

RESULT 8

US-08-481-968A-4
; Sequence 4, Application US/08481968A
; Patent No. 6300490
; GENERAL INFORMATION:
; APPLICANT: Huber, Brian
; TITLE OF INVENTION: Molecular Constructs Comprising a Carcinoembryonic Antigen (CEA)
; TITLE OF INVENTION: Transcriptional Regulatory Region
; FILE REFERENCE: PB1087054
; CURRENT APPLICATION NUMBER: US/08/481,968A
; CURRENT FILING DATE: 1998-06-07
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 11288
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-481-968A-4

Query Match 1.0%; Score 22; DB 4; Length 11288;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AAAAAAGAAAGAAAGAAAGA 92
|||||
DB 834 AAAAAAGAAAGAAAGAAAGA 855

RESULT 9

US-08-154-712B-4
; Sequence 4, Application US/08154712B
; Patent No. 6337209
; GENERAL INFORMATION:
; APPLICANT: Huber, Brian
; TITLE OF INVENTION: Molecular Constructs Containing a Carcinoembryonic Antigen Region
; TITLE OF INVENTION: Sequence
; FILE REFERENCE: PB1087053
; CURRENT APPLICATION NUMBER: US/08/154,712B
; CURRENT FILING DATE: 1993-11-19
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 11288
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-154-712B-4

Query Match 1.0%; Score 22; DB 4; Length 11288;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AAAAAAGAAAGAAAGAAAGA 92
|||||
DB 834 AAAAAAGAAAGAAAGAAAGA 855

RESULT 10
US-09-738-894A-3
; Sequence 3, Application US/09738894A
; Patent No. 6331423
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000636
; CURRENT APPLICATION NUMBER: US/09/738,894A
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36651
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(36651)
; OTHER INFORMATION: n = A,T,C or G
US-09-738-894A-3

Query Match 1.0%; Score 22; DB 4; Length 36651;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 AAAAAAGAAAGAAAGAAAGA 92
|||||
DB 5334 AAAAAAGAAAGAAAGAAAGA 5355

RESULT 11

US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H. CONTIG"
US-08-724-394A-20

Query Match 1.0%; Score 22; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 CAGCCTGGGTGACAGTGAGACT 49
|||||
DB 236746 CAGCCTGGGTGACAGTGAGACT 236767

RESULT 13
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H. CONTIG"
US-08-724-394A-22

Query Match 1.0%; Score 22; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 CAGCCTGGGTGACAGTGAGACT 49
|||||
DB 236746 CAGCCTGGGTGACAGTGAGACT 236767

RESULT 14
US-09-280-116-247
; Sequence 247, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:

SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H. CONTIG"
US-08-724-394A-20

Query Match 1.0%; Score 22; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 CAGCCTGGGTGACAGTGAGACT 49
|||||
DB 236746 CAGCCTGGGTGACAGTGAGACT 236767

RESULT 12
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H. CONTIG"
US-08-724-394A-21

Query Match 1.0%; Score 22; DB 2; Length 246240;

; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176955
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 247
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: hemoglobinase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(544)
; OTHER INFORMATION: n = a, t, c or g
US-09-280-116-247

Query Match 0.9%; Score 21; DB 4; Length 544;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

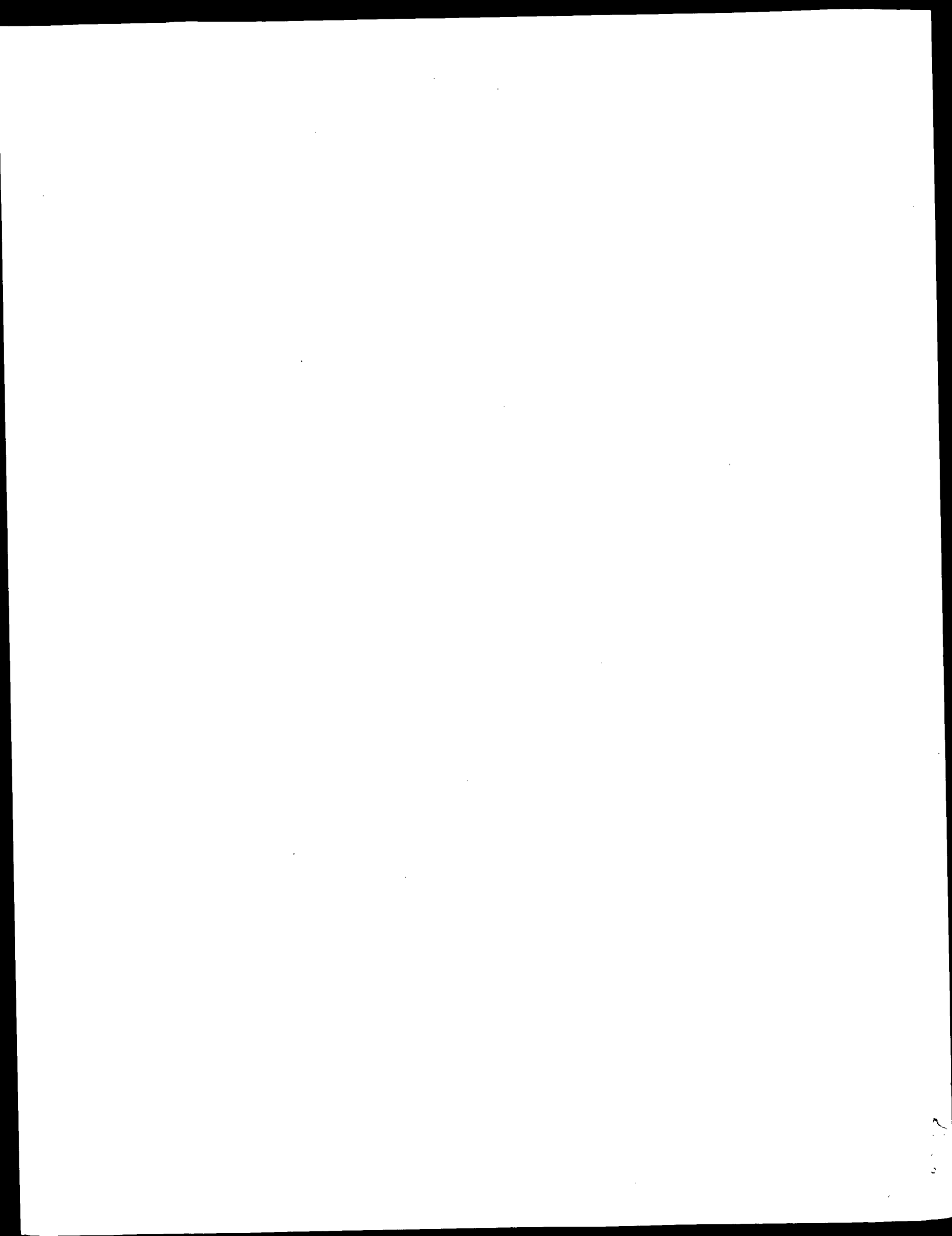
QY 41 AGTGAGACTCTGTCTCAACA 61
|||||
Db 370 AGTGAGACTCTGTCTCAACA 390

RESULT 15
US-09-328-111-68
; Sequence 68, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 68
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(579)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-68

Query Match 0.9%; Score 21; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2219 TTTATTAATAAAAAAAAAA 2239
|||||
Db 188 TTTATTAATAAAAAAAAAA 208

Search completed: April 16, 2003, 14:34:09
Job time : 1441 secs



QY	241	AGGCTGCTCCCTCCGCAACAGAGAGCCACCTGGCGCAGGAGCGCTTTTCAGGAAGAGAC	300
Db	241	AGGCTGCTCCCTCCGCAACAGAGAGCCACCTGGCGCAGGAGCGCTTTTCAGGAAGAGAC	300
QY	301	GCCTTTTCAGGAAGAGAGAGCCCTTTTCAGGAAGAGAGAGAGAGAGAGAGAGAGC	360
Db	301	GCCTTTTCAGGAAGAGAGAGAGCCCTTTTCAGGAAGAGAGAGAGAGAGAGAGAGAGC	360
QY	361	ACTTTACTGAGGAGC	420
Db	361	ACTTTACTGAGGAGC	420
QY	421	TCTCCTAAGGCGTGTCTCCAGAACAGCGGAGCGGTGGGAGATGCTCTGACCATCTGGACG	480
Db	421	TCTCCTAAGGCGTGTCTCCAGAACAGCGGAGCGGTGGGAGATGCTCTGACCATCTGGACG	480
QY	481	GTGTGGGGTCTGTCTGCTACTATTTGGAGCTTTGCTCTTATGCTGAATTTGGGAACACTATA	540
Db	481	GTGTGGGGTCTGTCTGCTACTATTTGGAGCTTTGCTCTTATGCTGAATTTGGGAACACTATA	540
QY	541	AAGAAATCTGGAGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTTACAGCTTTT	600
Db	541	AAGAAATCTGGAGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATTTACAGCTTTT	600
QY	601	GTACGAGTCTGGGTGGAACTCTCATTAACGCCCTGGAGCTACTGCTGTGATATCCCTG	660
Db	601	GTACGAGTCTGGGTGGAACTCTCATTAACGCCCTGGAGCTACTGCTGTGATATCCCTG	660
QY	661	GCATTTGGAGCCTACATTTCTGGAACCAATTTTATTTCAATGTGAATTTGGAACTGGG	720
Db	661	GCATTTGGAGCCTACATTTCTGGAACCAATTTTATTTCAATGTGAATTTGGAACTGGG	720
QY	721	ATCAAGCTCAATACAGCTGTGGGCAATTAACGTAGTGTGATGCTCTTAATAGCATGAGTCTC	780
Db	721	ATCAAGCTCAATACAGCTGTGGGCAATTAACGTAGTGTGATGCTCTTAATAGCATGAGTCTC	780
QY	781	AGCTGGAGCGCCGATCCAGATTTCTTAACCTTTTGAAGCTTCACAGCAATTTCTGATA	840
Db	781	AGCTGGAGCGCCGATCCAGATTTCTTAACCTTTTGAAGCTTCACAGCAATTTCTGATA	840
QY	841	ATTATAGTCCCTGGAGTTATGCACTAATTAAGGTCAACCGAGAGAACTTTAAGACGCC	900
Db	841	ATTATAGTCCCTGGAGTTATGCACTAATTAAGGTCAACCGAGAGAACTTTAAGACGCC	900
QY	901	TTTTTCAGGAAGAGATTCAGATTTACGCGTGTGCACTGGCTTTTATTTAGTGAATGTAT	960
Db	901	TTTTTCAGGAAGAGATTCAGATTTACGCGTGTGCACTGGCTTTTATTTAGTGAATGTAT	960
QY	961	GCATATGCTGGCTGTTTACCTCAACTTTGTTACTGAAGAGTAGAAAAACCTGAAAAA	1020
Db	961	GCATATGCTGGCTGTTTACCTCAACTTTGTTACTGAAGAGTAGAAAAACCTGAAAAA	1020
QY	1021	ACATTTCCCTTTGCAATATGATATCCATGGCATTTGCTACCATTTGCTGTGCTGACA	1080
Db	1021	ACATTTCCCTTTGCAATATGATATCCATGGCATTTGCTACCATTTGCTGTGCTGACA	1080
QY	1081	AATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAAGTGGCA	1140
Db	1081	AATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAAGTGGCA	1140
QY	1141	GTGACCTTTTTCAGGCGTACTGGGAAATTTCTCATTTAGCAGTTCCGATCTTTGTTGCC	1200
Db	1141	GTGACCTTTTTCAGGCGTACTGGGAAATTTCTCATTTAGCAGTTCCGATCTTTGTTGCC	1200
QY	1201	CTCTCTGCTTTTGGCTCCATGAACGCTGTGTGTTTGGTCTGCTCCAGGTTATTTCTATGTT	1260
Db	1201	CTCTCTGCTTTTGGCTCCATGAACGCTGTGTGTTTGGTCTGCTCCAGGTTATTTCTATGTT	1260
QY	1261	GCCTCTGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCACACT	1320
Db	1261	GCCTCTGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCACACT	1320
QY	1321	CCTCTACAGCTGTTATTTTTCACACCCCTTTGACAAATGATGCTCTCTCTGGAGAC	1380
Db	1321	CCTCTACAGCTGTTATTTTTCACACCCCTTTGACAAATGATGCTCTCTCTGGAGAC	1380
QY	1381	CTCGACAGTCTTTTTCAGATTTTCTCAGTTTTCAGAGTGGCTTTTATTTAGGCTGGCAGTT	1440
Db	1381	CTCGACAGTCTTTTTCAGATTTTCTCAGTTTTCAGAGTGGCTTTTATTTAGGCTGGCAGTT	1440
QY	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATGCTCTCTTTCAAGGTGCCA	1500
Db	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATGCTCTCTTTCAAGGTGCCA	1500
QY	1501	CTGTTCATCCAGCTTTTGTTCCTTCACATGCCCTCTTCATGCTTTCCTCTCTAT	1560
Db	1501	CTGTTCATCCAGCTTTTGTTCCTTCACATGCCCTCTTCATGCTTTCCTCTCTAT	1560
QY	1561	TGGGACCCATTTTAGTACAGGGAATGCTTTCGCTCATCTCTGACTGGAGTCCCTCGSTAT	1620
Db	1561	TGGGACCCATTTTAGTACAGGGAATGCTTTCGCTCATCTCTGACTGGAGTCCCTCGSTAT	1620
QY	1621	TATCTCTTTATTTATGGGACAAAGAACCCAGGTGGTGTAGAAATGCTCAGAGAAATA	1680
Db	1621	TATCTCTTTATTTATGGGACAAAGAACCCAGGTGGTGTAGAAATGCTCAGAGAAATA	1680
QY	1681	ACCAGAACATTTACAAATTAATGCTGGAAGTGTACAGAGAGAGATAAGTATGAACATAA	1740
Db	1681	ACCAGAACATTTACAAATTAATGCTGGAAGTGTACAGAGAGAGATAAGTATGAACATAA	1740
QY	1741	GGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAT	1800
Db	1741	GGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAT	1800
QY	1801	TTTCTGAAAGCTCTAGAGAAATTAACAATTTGGTGTGATTAACAAAAGGAGTCAATTTTAA	1860
Db	1801	TTTCTGAAAGCTCTAGAGAAATTAACAATTTGGTGTGATTAACAAAAGGAGTCAATTTTAA	1860
QY	1861	TTTCTGAAAGCTCTAGAGAAATTAACAATTTGGTGTGATTAACAAAAGGAGTCAATTTTAA	1920
Db	1861	TTTCTGAAAGCTCTAGAGAAATTAACAATTTGGTGTGATTAACAAAAGGAGTCAATTTTAA	1920
QY	1921	TTATAGAAAGTGAATATGCAATTTATGCTGAGTCCGCAATTTGAGTCTCTGATACC	1980
Db	1921	TTATAGAAAGTGAATATGCAATTTATGCTGAGTCCGCAATTTGAGTCTCTGATACC	1980
QY	1981	TACCTATTTGGGTTAGGAGAAAGAGCTAGCAATTTACTATGCTGCTCATTTCTTACAACAT	2040
Db	1981	TACCTATTTGGGTTAGGAGAAAGAGCTAGCAATTTACTATGCTGCTCATTTCTTACAACAT	2040
QY	2041	ATGTTAGCACGGCAAGAAACCTTCAAAATTTGAAGCTGAGATTTTCTGTATATATGGGTT	2100
Db	2041	ATGTTAGCACGGCAAGAAACCTTCAAAATTTGAAGCTGAGATTTTCTGTATATATGGGTT	2100
QY	2101	TTGTTAAAGATGGTTTACACACTACAGATGCTATCTGTGAAAAGTGTTCATTTCTG	2160
Db	2101	TTGTTAAAGATGGTTTACACACTACAGATGCTATCTGTGAAAAGTGTTCATTTCTG	2160
QY	2161	AAAAAAGCATACATCATGATTATGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2220
Db	2161	AAAAAAGCATACATCATGATTATGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2220
QY	2221	TATTTAAAAAAG	2280
Db	2221	TATTTAAAAAAG	2280

RESULT 2
 US-09-738-973-440
 ; Sequence 440, Application US/09738973
 ; Patent No. US2002011056A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Lodes, Michael J.

; APPLICANT: Fling, Steven P.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Secret, Heather
 ; APPLICANT: Indrias, Carol Yoseph
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Elliot, Mark
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Kalos, Michael D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; FILE REFERENCE: 210121.475C9
 ; CURRENT APPLICATION NUMBER: US/09/738, 973
 ; CURRENT FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 587
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 440
 ; LENGTH: 2239
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-738-973-440

Query Match 100.0%; Score 2239; DB 10; Length 2239;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGAGTTGAGTGAGCAGAGATCATGCCCTGGGTGACAGTGAGACTCTGCTCAAC	60
Db	1	GGAGTTGAGTGAGCAGAGATCATGCCCTGGGTGACAGTGAGACTCTGCTCAAC	60
Qy	61	AGAAATTAAGGAAAAAGAAAGAAAGAGAGAGAGAGAAATCCAGGCCAATGTGG	120
Db	61	AGAAATTAAGGAAAAAGAAAGAAAGAGAGAGAGAGAAATCCAGGCCAATGTGG	120
Qy	121	CATAGATTTATCATATCTGGATTTTGGATTTCTTTTGTCTCATCTGATCA	180
Db	121	CATAGATTTATCATATCTGGATTTTGGATTTCTTTTGTCTCATCTGATCA	180
Qy	181	GGAAGGCTCTGTGTCCACATCTCCAAAGGAGTTACTCGAGGGAATTTAACGG	240
Db	181	GGAAGGCTCTGTGTCCACATCTCCAAAGGAGTTACTCGAGGGAATTTAACGG	240
Qy	241	AGGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCGCTTTTCAGGAGAGAC	300
Db	241	AGGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCGCTTTTCAGGAGAGAC	300
Qy	301	GCCTTTTCAGGAGAGACGCCCTTTTCAGGAGAGAGAAAGTGCAGCTGAAGAGAAATC	360
Db	301	GCCTTTTCAGGAGAGAGACGCCCTTTTCAGGAGAGAGAAAGTGCAGCTGAAGAGAAATC	360
Qy	361	ACTTTACTGAGGAGGAGTCTCCATTTATTCATTTGGCACCACCATTTGGAGAGAGATCTTCATC	420
Db	361	ACTTTACTGAGGAGGAGTCTCCATTTATTCATTTGGCACCACCATTTGGAGAGAGATCTTCATC	420
Qy	421	TCCTCTAAGGGGCTGTGCAGAACAGGGCAGCGTGGGAGTCTCTGACCATCTGGAGC	480
Db	421	TCCTCTAAGGGGCTGTGCAGAACAGGGCAGCGTGGGAGTCTCTGACCATCTGGAGC	480
Qy	481	GTGTGTGGGCTCTGTCTACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACTATA	540
Db	481	GTGTGTGGGCTCTGTCTACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACTATA	540
Qy	541	AGAAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTGGTCCATTACGAGCTTTT	600
Db	541	AGAAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTGGTCCATTACGAGCTTTT	600
Qy	601	GTACAGTCTGGTGGAACTCTCATATATACGCCCTGCAGTACTGCTGTATCCCTG	660
Db	601	GTACAGTCTGGTGGAACTCTCATATATACGCCCTGCAGTACTGCTGTATCCCTG	660
Qy	661	GCATTTGGAGCTACATTTCTGGAACCATTTTATTTATCAATCTGAATCCCTGACCTTGG	720
Db	661	GCATTTGGAGCTACATTTCTGGAACCATTTTATTTATCAATCTGAATCCCTGACCTTGG	720

Qy	721	ATCAAGCTCATTACAGCTGTGGCATACCTAGTGTGCTCTTAATAGCATGAGTGT	780
Db	721	ATCAAGCTCATTACAGCTGTGGCATAACTAGTGTGCTCTTAATAGCATGAGTGT	780
Qy	781	AGCTGGAGGCCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATA	840
Db	781	AGCTGGAGGCCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATA	840
Qy	841	ATTATAGTCCCTGGAGTATTCAGCTAATTAAGGTCAAACGAGCACTTTAAAGAGCC	900
Db	841	ATTATAGTCCCTGGAGTATTCAGCTAATTAAGGTCAAACGAGCACTTTAAAGAGCC	900
Qy	901	TTTTAGGAAGAGATTTCAAGTATTACCGGTGGCAGTGGCTTTTATTTATGAAGTAT	960
Db	901	TTTTAGGAAGAGATTTCAAGTATTACCGGTGGCAGTGGCTTTTATTTATGAAGTAT	960
Qy	961	GCATATGCTGCTGGTTTACCTCAACTTTGTTACTGAGAGTAGAAAAACCCGAAAA	1020
Db	961	GCATATGCTGCTGGTTTACCTCAACTTTGTTACTGAGAGTAGAAAAACCCGAAAA	1020
Qy	1021	ACCATCCCTTGCATATGATATCCATGGCCATTTGTCACCATTTGGCTATGTGTGACA	1080
Db	1021	ACCATCCCTTGCATATGATATCCATGGCCATTTGTCACCATTTGGCTATGTGTGACA	1080
Qy	1081	AATGTGCTTACTTTAGCACCATTAATGCTGAGGAGTCTGCTTTCAAAATGCAGTGCA	1140
Db	1081	AATGTGCTTACTTTAGCACCATTAATGCTGAGGAGTCTGCTTTCAAAATGCAGTGCA	1140
Qy	1141	GTGACCTTTTCTGAGGGCTACTGGGAAATTTCTCATAGGAGTCCGATCTTTGTGGC	1200
Db	1141	GTGACCTTTTCTGAGGGCTACTGGGAAATTTCTCATAGGAGTCCGATCTTTGTGGC	1200
Qy	1201	CTCTCTGCTTTGGCTCCATGAAACGGTGGTGTGTTGCTGTCTCCAGTATTTCTATGTT	1260
Db	1201	CTCTCTGCTTTGGCTCCATGAAACGGTGGTGTGTTGCTGTCTCCAGTATTTCTATGTT	1260
Qy	1261	GCCTCTGAGAGGCTCACCTTCCAGAAATCTCCATGATTCATGTCGGGAGGACACT	1320
Db	1261	GCCTCTGAGAGGCTCACCTTCCAGAAATCTCCATGATTCATGTCGGGAGGACACT	1320
Qy	1321	CTCTTACAGCTGTATTTGTTTGCACCTTTGACAATGATATCTCTCTCGAGAC	1380
Db	1321	CTCTTACAGCTGTATTTGTTTGCACCTTTGACAATGATATCTCTCTCGAGAC	1380
Qy	1381	CTGAGAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTTGGGCTGGCAGTT	1440
Db	1381	CTGAGAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTTGGGCTGGCAGTT	1440
Qy	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGCTTCAAGGTGCCA	1500
Db	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGCTTCAAGGTGCCA	1500
Qy	1501	CTGTTTATCCAGCTTTGTTTTCATACATGCTCTTCATGGTGGCCTTTCCCTCTAT	1560
Db	1501	CTGTTTATCCAGCTTTGTTTTCATACATGCTCTTCATGGTGGCCTTTCCCTCTAT	1560
Qy	1561	TCGGACCCATTTAGTACAGGATTTGGCTTGGTCACTCTGAGTGGGCTGGCTAT	1620
Db	1561	TCGGACCCATTTAGTACAGGATTTGGCTTGGTCACTCTGAGTGGGCTGGCTAT	1620
Qy	1621	TATCTCTTTATATATGGGCAAGAAACCCAGGTGGTTAGAAATATGTCAGAGAAAAA	1680
Db	1621	TATCTCTTTATATATGGGCAAGAAACCCAGGTGGTTAGAAATATGTCAGAGAAAAA	1680
Qy	1681	ACCAGAACATTAATAATGCTGAAAGTGTACAGAGAAAGATAAGTTATGAAGTAT	1740
Db	1681	ACCAGAACATTAATAATGCTGAAAGTGTACAGAGAAAGATAAGTTATGAAGTAT	1740
Qy	1741	GSACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGATTTTACTTCAT	1800
Db	1741	GSACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGATTTTACTTCAT	1800

QY 1801 TTTCTGAAAGCTAGAGAAATACAACTTTGGTGATATAAACAAGAGAGTCAAGTATATTTTA 1860
Db 1801 TTTCTGAAAGCTAGAGAAATACAACTTTGGTGATATAAACAAGAGAGTCAAGTATATTTTA 1860
QY 1861 TTCATATATTTTACCATATTCGAACATTAATTTCTAAGAAATTTAGTTATACCTCTATGTAG 1920
Db 1861 TTCATATATTTTACCATATTCGAACATTAATTTCTAAGAAATTTAGTTATACCTCTATGTAG 1920
QY 1921 TTATAGAAAGCTAGATATGCAATTTCTATGAGTCGACAAATTTCTGAGTCTCTGATACC 1980
Db 1921 TTATAGAAAGCTAGATATGCAATTTCTATGAGTCGACAAATTTCTGAGTCTCTGATACC 1980
QY 1981 TACCTATTTGGGTTAGGAGAAAGACATAGACAAATTTATGAGTCTCTACACAT 2040
Db 1981 TACCTATTTGGGTTAGGAGAAAGACATAGACAAATTTATGAGTCTCTACACAT 2040
QY 2041 ATGTTAGCAGCGCAAGAACCTTCAAAATGGAAGTCTGAGATTTTCTGTATATATGGTT 2100
Db 2041 ATGTTAGCAGCGCAAGAACCTTCAAAATGGAAGTCTGAGATTTTCTGTATATATGGTT 2100
QY 2101 TTGTAAGATGTTTACACACTACAGATGCTATACATGCTGAAAGTGTTCATTTCTG 2160
Db 2101 TTGTAAGATGTTTACACACTACAGATGCTATACATGCTGAAAGTGTTCATTTCTG 2160
QY 2161 AAAAAAGCATCATCATGATATGGAAGAGGAGAGAGTGTGTTCTTAAAT 2220
Db 2161 AAAAAAGCATCATCATGATATGGAAGAGGAGAGAGTGTGTTCTTAAAT 2220
QY 2221 TATTAATAAAAAAAAAA 2239
Db 2221 TATTAATAAAAAAAAAA 2239

RESULT 3
US-09-854-133-441
; Sequence 441, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Iodes, Michael J.
; APPLICANT: Mohamath, Raodoh.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secretist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 441
; LENGTH: 5981
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-854-133-441

Query Match 83.4%; Score 1867; DB 9; Length 5981;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTTATCATTTGGC 393
Db 278 GAGAAAGTGCAGCTGGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTTATCATTTGC 337
QY 394 ACCATCATTTGAGCAGAGAAATTTCTATCTCCCTAAGGGCGTGTCCAGAACACAGGGCAGC 453
Db 338 ACCATCATTTGAGCAGAGAAATTTCTATCTCTAAGGGCGTGTCCAGAACACAGGGCAGC 397
QY 454 GTGGGCGATGCTCTGACCATCTTGACCGGTGTGTGGGGTCTCTGACATTTTGGAGCTTTG 513
Db 398 GTGGGCGATGCTCTGACCATCTTGACCGGTGTGTGGGGTCTCTGACATTTTGGAGCTTTG 457
QY 514 TCTTATGCTGAATTTGGGACAACTATAAAGAAATCTGGAGTCTATACATATATTTTG 573

Db 458 TCTTATGCTGAATTTGGGAAACAATATAAAGAAATCTGGAGGTCATTTACACATATATTTTG 517
QY 574 GAAGTCTTTGGTCCATTTACCAGCTTTTCTAGAGTCTGGTGGAACTCTCTCATATATACGC 633
Db 518 GAAGTCTTTGGTCCATTTACCAGCTTTTCTAGAGTCTGGTGGAACTCTCTCATATATACGC 577
QY 634 CCTGAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCAATTTTTT 693
Db 578 CCTGAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCAATTTTTT 637
QY 694 ATTCAATGTGAATCCCTGAACTTGGATCAAGCTCATTTACAGCTGTGGGATCAACTGTA 753
Db 638 ATTCAATGTGAATCCCTGAACTTGGATCAAGCTCATTTACAGCTGTGGGATCAACTGTA 697
QY 754 GTGATGGTCTTAATAGCATGAGTGTGAGTGGAGCGCCGATCCAGATTTTCTTAACC 813
Db 698 GTGATGGTCTTAATAGCATGAGTGTGAGTGGAGCGCCGATCCAGATTTTCTTAACC 757
QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGACGATTAATAAA 873
Db 758 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGACGATTAATAAA 817
QY 874 GGTCAACGCGAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTTACGCGGTTG 933
Db 818 GGTCAACGCGAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTTACGCGGTTG 877
QY 934 CCACTGGCTTTTATGGAATGTATGATATGCTGGCTGTTTACCTCAACTTTGTT 993
Db 878 CCACTGGCTTTTATGGAATGTATGATATGCTGGCTGTTTACCTCAACTTTGTT 937
QY 994 ACTGAAGAGTGAAGAACCTGAAAAACCAATCCCTTGCATATATGATATCCATGGCC 1053
Db 938 ACTGAAGAGTGAAGAACCTGAAAAACCAATCCCTTGCATATATGATATCCATGGCC 997
QY 1054 ATTGTCAACATTTGCTATGTGTCACAAATGTGGCTACTTTACGACCACTTAATGCTGAG 1113
Db 998 ATTGTCAACATTTGCTATGTGTCACAAATGTGGCTACTTTACGACCACTTAATGCTGAG 1057
QY 1114 GAGCTGCTGCTTTCAAAATGCAAGTGGCAGTGAACCTTTCTGAGCGGCTACTGGGAAATTTTC 1173
Db 1058 GAGCTGCTGCTTTCAAAATGCAAGTGGCAGTGAACCTTTCTGAGCGGCTACTGGGAAATTTTC 1117
QY 1174 TCATTAGCAGTTCGAGTCTTTGTTGGCTCTCTGCTTTGGCTCCATGAAAGTGTGTG 1233
Db 1118 TCATTAGCAGTTCGAGTCTTTGTTGGCTCTCTGCTTTGGCTCCATGAAAGTGTGTG 1177
QY 1234 TTTGCTGCTCCAGGTTATTTATGTTGGCTGTGAGAGGGTCACTTCCAGAAATCCCTC 1293
Db 1178 TTTGCTGCTCCAGGTTATTTATGTTGGCTGTGAGAGGGTCACTTCCAGAAATCCCTC 1237
QY 1294 TCCATGATTCAATGTCGCAAGCAGCAGTCTCTACAGCTGTTATGTTTTCACCCCTTTG 1353
Db 1238 TCCATGATTCAATGTCGCAAGCAGCAGTCTCTACAGCTGTTATGTTTTCACCCCTTTG 1297
QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCC 1413
Db 1298 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCC 1357
QY 1414 AGGTGGCTTTTATTTGGCTGGCAGTGTGCTGGCTGATTTATCTTCGATACAAATGGCCA 1473
Db 1358 AGGTGGCTTTTATTTGGCTGGCAGTGTGCTGGCTGATTTATCTTCGATACAAATGGCCA 1417
QY 1474 GATATGCAATGCTCTCTCAAGGTGCCACTGTTTCAATCCAGCTTTGTTTCTTCCATATG 1533
Db 1418 GATATGCAATGCTCTCTCAAGGTGCCACTGTTTCAATCCAGCTTTGTTTCTTCCATATG 1477
QY 1534 CTCCTTCATGTTGGCTTTCCCTCTATTTGGACCCATTTAGTACAGGATTTGGCTTGGTC 1593
Db 1478 CTCCTTCATGTTGGCTTTCCCTCTATTTGGACCCATTTAGTACAGGATTTGGCTTGGTC 1537
QY 1594 ATCACTCTGACGTGAGTCCCTGCGTATTTATCTTTATTTATATGGACAAAGAAACCCAGG 1653

Db 1538 ATCACTCTGACTGGAGTCCCTGCGTATATCTCTTTATTATATGGGCAAGAACCCAGG 1597
QY 1654 TGGTTTAGAATAATCTCAGAGAAATAACCAACATTTACAAATTAATCTGGAAGTTGTA 1713
Db 1598 TGGTTTAGAATAATCTCAGAGAAATAACCAACATTTACAAATTAATCTGGAAGTTGTA 1657
QY 1714 CCAGAAGAAGATAAGTTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGA 1773
Db 1658 CCAGAAGAAGATAAGTTTATGAACCTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGGA 1717
QY 1774 GACCAAAATAGGAGTTTACTTCAATTTCTGAAAGCTTAGAGAAATTAACAATTTGGTG 1833
Db 1718 GACCAAAATAGGAGTTTACTTCAATTTCTGAAAGCTTAGAGAAATTAACAATTTGGTG 1777
QY 1834 ATAAACAAAGAGTCAAGTTTATTTTATTCATATATTTTAGCATATTTGCAACTAATTTCT 1893
Db 1778 ATAAACAAAGAGTCAAGTTTATTTTATTCATATATTTTAGCATATTTGCAACTAATTTCT 1837
QY 1894 AAGAAATTTAGTTTAACTCTATAGTTATAGAAAGTGAATATGCAAGTTTATTTCTATGAG 1953
Db 1838 AAGAAATTTAGTTTAACTCTATAGTTATAGAAAGTGAATATGCAAGTTTATTTCTATGAG 1897
QY 1954 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGGTTAGGAGAAAGACTAGACAA 2013
Db 1898 TCGCACAAATCTTGAGTCTCTGATACCTACCTATTTGGGGTTAGGAGAAAGACTAGACAA 1957
QY 2014 TTACTATGTGCTATCTCTACAAACATATGTTAGCAGCGCAAGAAACCTTCAAAATGAAG 2073
Db 1958 TTACTATGTGCTATCTCTACAAACATATGTTAGCAGCGCAAGAAACCTTCAAAATGAAG 2017
QY 2074 ACTGAGATTTTCTGATATATGTTGTTTGTAAAGATGTTTACACACTACAGATGTCT 2133
Db 2018 ACTGAGATTTTCTGATATATGTTGTTTGTAAAGATGTTTACACACTACAGATGTCT 2077
QY 2134 ATACTGTGAAAGTGTCTTCAATTTCTGAAAGAAAGCATAACATCATGATTATGGCAAAG 2193
Db 2078 ATACTGTGAAAGTGTCTTCAATTTCTGAAAGAAAGCATAACATCATGATTATGGCAAAG 2137
QY 2194 GAGAGAA 2200
Db 2138 GAGAGAA 2144

RESULT 4
US-09-738-973-441
; Sequence 441, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secretist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 441
; LENGTH: 5981
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-973-441

Query Match 83.4%; Score 1867; DB 10; Length 5981;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATATCATTTGCG 393
Db 278 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATATCATTTGCG 337
QY 394 ACCATCATTTGGAGCAGGAATCTTTCATCTCTCTTAAGGGCTGTCTCAGAACACGGGCGAG 453
Db 338 ACCATCATTTGGAGCAGGAATCTTTCATCTCTCTTAAGGGCTGTCTCAGAACACGGGCGAG 397
QY 454 GTGGGCTATGCTCTGACCATCTGGAGCGGTGTGTGGGGTCTGTGCACTATTTGGAGCTTTG 513
Db 398 GTGGGCTATGCTCTGACCATCTGGAGCGGTGTGTGGGGTCTGTGCACTATTTGGAGCTTTG 457
QY 514 TCTTATGCTGAATTTGGGAAACAATATAAGAAATCTGAGGTCATTACACATATATTTTG 573
Db 458 TCTTATGCTGAATTTGGGAAACAATATAAGAAATCTGAGGTCATTACACATATATTTTG 517
QY 574 GAAGTCTTTGGTCCATTACAGCTTTGTACGAGTCTGGGTGGAACTCCTCATATACGCG 633
Db 518 GAAGTCTTTGGTCCATTACAGCTTTGTACGAGTCTGGGTGGAACTCCTCATATACGCG 577
QY 634 CCGTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 693
Db 578 CCGTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 637
QY 694 ATTCAATGTGAATCCCTGAACCTTGGATCAAGCTCATTACAGCTGTGGGCAATACTGTA 753
Db 638 ATTCAATGTGAATCCCTGAACCTTGGATCAAGCTCATTACAGCTGTGGGCAATACTGTA 697
QY 754 GTGATGGTCTTAATAGCATGATGTGTCAGTGGAGCGCCGGATCCAGATTTCTTAAACC 813
Db 698 GTGATGGTCTTAATAGCATGATGTGTCAGTGGAGCGCCGGATCCAGATTTCTTAAACC 757
QY 814 TTTTGCAGCTCACAGCAATTTCTGATAATATTAGTCCCTGGAGTTATGAGCAATAATTA 873
Db 758 TTTTGCAGCTCACAGCAATTTCTGATAATATTAGTCCCTGGAGTTATGAGCAATAATTA 817
QY 874 GGTCAAGCGCAGAACTTTAAAGACGCTTTTAAAGACGAGATTCAAGTATTACGGGTTG 933
Db 818 GGTCAAGCGCAGAACTTTAAAGACGCTTTTAAAGACGAGATTCAAGTATTACGGGTTG 877
QY 934 CCAGTGGCTTTTATATGGAATGATGATGATGCTGGTGTGTTTACCTCAACTTTGTT 993
Db 878 CCAGTGGCTTTTATATGGAATGATGATGATGCTGGTGTGTTTACCTCAACTTTGTT 937
QY 994 ACTGAAGAAGTAGAAAAACCCCTGAAAAAACCATTCCTCCCTTGCATATGATATCATGCC 1053
Db 938 ACTGAAGAAGTAGAAAAACCCCTGAAAAAACCATTCCTCCCTTGCATATGATATCATGCC 997
QY 1054 ATTGTCAACATTTGGCTATGCTGTGACAAATGTGGCCTACTTTAGGACCATTAATGCTGAG 1113
Db 998 ATTGTCAACATTTGGCTATGCTGTGACAAATGTGGCCTACTTTAGGACCATTAATGCTGAG 1057
QY 1114 GAGCTGCTGCTTCAATGCGAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 1058 GAGCTGCTGCTTCAATGCGAGTGGCAGTGACCTTTCTGAGCGGCTACTGGGAAATTC 1117
QY 1174 TCATTAGCAGTTCCGATCTTTGTCCTCTCTGCTGCTTGGCTCCATGACGAGGTTG 1233
Db 1118 TCATTAGCAGTTCCGATCTTTGTCCTCTCTGCTGCTTGGCTCCATGACGAGGTTG 1177
QY 1234 TTTGCTGTCTCCAGTTATTTCTATGTTGCGTCTCGAGAGGCTACCTTCCAGAAATCCT 1293
Db 1178 TTTGCTGTCTCCAGTTATTTCTATGTTGCGTCTCGAGAGGCTACCTTCCAGAAATCCT 1237
QY 1294 TCCATGATTTCATGTCGCAAGCACACCTCTACAGCTGTTATGTTGTCACCTTTG 1353
Db 1238 TCCATGATTTCATGTCGCAAGCACACCTCTACAGCTGTTATGTTGTCACCTTTG 1297
QY 1354 ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGCTCTTTTGAATTTTCCCTCAGTTTGGC 1413

Db	1298	ACAATGATATGCTCTCTCTCTCGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGTGC	1357
QY	1414	AGTGGCTTTTATTTGGCTGGCAGTTGCTGGCTGATTATCTTCGATACAAATGCCCA	1473
Db	1358	AGTGGCTTTTATTTGGCTGGCAGTTGCTGGCTGATTATCTTCGATACAAATGCCCA	1417
QY	1474	GATATGCATCGCTTTTCAAGGTGCGACATGTTCAATCCAGCTTGTGTTTCCCTTCACATGC	1533
Db	1418	GATATGCATCGCTTTTCAAGGTGCGACATGTTCAATCCAGCTTGTGTTTCCCTTCACATGC	1477
QY	1534	CTCTTCATGTTGCCCTTTCCCTCTATTCCGACCCATTATAGTACAGGGATTGGCTTCGTC	1593
Db	1478	CTCTTCATGTTGCCCTTTCCCTCTATTCCGACCCATTATAGTACAGGGATTGGCTTCGTC	1537
QY	1594	ATCACCTCTGACTGGAGTCCCTCGGTATATATCTTTTATATATGGGACAAGAAACCCAGG	1653
Db	1538	ATCACCTCTGACTGGAGTCCCTCGGTATATATCTTTTATATATGGGACAAGAAACCCAGG	1597
QY	1654	TGGTTTAGAATAATGTTCAGAGAAAATAACAGAAACATTACAAATAATCTGGAAGTTGTA	1713
Db	1598	TGGTTTAGAATAATGTTCAGAGAAAATAACAGAAACATTACAAATAATCTGGAAGTTGTA	1657
QY	1714	CCAGAGAAGATAAGTTATGAACATAATGGACTTGGAGATCTGGCAATCTGCCCAAGGGGA	1773
Db	1658	CCAGAGAAGATAAGTTATGAACATAATGGACTTGGAGATCTGGCAATCTGCCCAAGGGGA	1717
QY	1774	GACACAAATAGGGATTTTACTTTCTCAAAAGTCTAGAGAATTACAACTTTGGTG	1833
Db	1718	GACACAAATAGGGATTTTACTTTCTCAAAAGTCTAGAGAATTACAACTTTGGTG	1777
QY	1834	ATAACAAAAGGAGTCAAGTTATTTTATTCATATATTTTAGCATATTCGAACCTAATTCT	1893
Db	1778	ATAACAAAAGGAGTCAAGTTATTTTATTCATATATTTTAGCATATTCGAACCTAATTCT	1837
QY	1894	AAGAAATTTAGTTATACTCTATGTAGTTATAGAAAGTGAATATGCAGTTATTCATGAG	1953
Db	1838	AAGAAATTTAGTTATACTCTATGTAGTTATAGAAAGTGAATATGCAGTTATTCATGAG	1897
QY	1954	TCGCACAAATTCCTGAGTCTCTGATACCTACTATTTGGGTTTAGGAGAAAGACTAGACAA	2013
Db	1898	TCGCACAAATTCCTGAGTCTCTGATACCTACTATTTGGGTTTAGGAGAAAGACTAGACAA	1957
QY	2014	TTACTATGTGGTCAATCTCTACAACATATGTTAGCAGCGCAAGAACCTTCAAATTTGAAG	2073
Db	1958	TTACTATGTGGTCAATCTCTACAACATATGTTAGCAGCGCAAGAACCTTCAAATTTGAAG	2017
QY	2074	ACTGAGATTTTCTGTATATATGCGTTTGTAAAGATGGTTTTACACACTACAGATGCTCT	2133
Db	2018	ACTGAGATTTTCTGTATATATGCGTTTGTAAAGATGGTTTTACACACTACAGATGCTCT	2077
QY	2134	ATACTGTCAAAAGTGTTTTCAATTTCTGAAAAAAGCATACATCATGATTTATGGCAAAAG	2193
Db	2078	ATACTGTCAAAAGTGTTTTCAATTTCTGAAAAAAGCATACATCATGATTTATGGCAAAAG	2137
QY	2194	GAGAGAA 2200	
Db	2138	GAGAGAA 2144	

RESULT 5
US-10-163-866-31
; Sequence 31, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605

; PRIOR FILING DATE: 2001-10-10	
; PRIOR APPLICATION NUMBER: US 60/338,733	
; PRIOR FILING DATE: 2001-10-22	
; PRIOR APPLICATION NUMBER: US 60/357,253	
; PRIOR FILING DATE: 2002-02-15	
; PRIOR APPLICATION NUMBER: US 60/357,600	
; PRIOR FILING DATE: 2002-02-15	
; NUMBER OF SEQ ID NOS: 54	
; SOFTWARE: PatentIn version 3.1	
; SEQ ID NO 31	
; LENGTH: 3144	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-10-163-866-31	
Query Match 81.1%; Score 1816; DB 9; Length 3144;	
Best Local Similarity 99.9%; Pred. No. 0;	
Matches 1866; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	334 GAGAAAGTCGACCTGAAGAGGAAAGTCACATTTACTGAGGAGAGTCCTCCATTAATCATTTGGC 393
Dd	106 GAGAAAGTCGACCTGAAGAGGAAAGTCACATTTACTGAGGAGAGTCCTCCATTAATCATTTGGC 165
Qy	394 ACCATATTGGAGCAGGAATCTTCATCTCCTAAGGCGTGCTCCAGAACACGGGCACG 453
Dd	166 ACCATATTGGAGCAGGAATCTTCATCTCCTAAGGCGTGCTCCAGAACACGGGCACG 225
Qy	454 GTGGGCATGCTCTGACCATCTGACCGTGCTGGGGTCTGTCTACATTAATTTGGAGCTTTG 513
Dd	226 GTGGGCATGCTCTGACCATCTGACCGTGCTGGGGTCTGTCTACATTAATTTGGAGCTTTG 285
Qy	514 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGTCAATACACATATATTTTG 573
Dd	286 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGTCAATACACATATATTTTG 345
Qy	574 GAAGTCTTTGGTCCATTACACGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGC 633
Dd	346 GAAGTCTTTGGTCCATTACACGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGC 405
Qy	634 CTTGAGCTACTGCTGTGATATCCCTGGCATTTGGAGCTACATCTCGAACCATTTTTT 693
Dd	406 CTTGAGCTACTGCTGTGATATCCCTGGCATTTGGAGCTACATCTCGAACCATTTTTT 465
Qy	694 ATTCAATGTGAATCCCTGAACTTCGATCAAGCTCATTACAGCTGTGGGCATAACTGTA 753
Dd	466 ATTCAATGTGAATCCCTGAACTTCGATCAAGCTCATTACAGCTGTGGGCATAACTGTA 525
Qy	754 GTGATGCTCCTAAATAGCATGATGTACGCTGGAGCGCCCGATCCAGATTTCTTTAACC 813
Dd	526 GTGATGCTCCTAAATAGCATGATGTACGCTGGAGCGCCCGATCCAGATTTCTTTAACC 595
Qy	814 TTTTGAAGCTCAGCAATTCGTATAATTATAGTCCCTGGAGTTATGACGCTAAATTTAA 873
Dd	586 TTTTGAAGCTCAGCAATTCGTATAATTATAGTCCCTGGAGTTATGACGCTAAATTTAA 645
Qy	874 GGTCAAAACGAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAGTATTACCGGTTG 933
Dd	646 GGTCAAAACGAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAGTATTACCGGTTG 705
Qy	934 CCAATGCGTTTTTATTATGGAATGATGCAATATGCTGGCTGGTTTACCTCAACTTTGTT 993
Dd	706 CCAATGCGTTTTTATTATGGAATGATGCAATATGCTGGCTGGTTTACCTCAACTTTGTT 765
Qy	994 ACTCAAGAAGCTAGAAAACCTGAAAAACCATTCCTCCTTGAATATGTATATCCATGGCC 1053
Dd	766 ACTGAAGAAGCTAGAAAACCTGAAAAACCATTCCTCCTGCAATATGTATATCCATGGCC 825
Qy	1054 ATTTGTCACCAATTTGGCTATGTGCTGACAAATGTGGGCTACTTTACGACCAATTAATGCTGAG 1113
Dd	826 ATTTGTCACCAATTTGGCTATGTGCTGACAAATGTGGGCTACTTTACGACCAATTAATGCTGAG 885
Qy	1114 GAGCTGCTGCTTTCAAATGCAGTGGCAGTGACCTTTTCTTGACGCGCTACTGGGAAATTC 1173

Db 886 GAGCTGCTGCTTCAAAGCACTGGCAGTGACCTTTTCTGAGCGGTACTGGGAAATTC 945

QY 1174 TCATTAGCAGTTCGCACTTTTGTGCTCCCTCTCTGCTTTGGCTCCATCAACGGTGGTGTG 1233

Db 946 TCATTAGCAGTTCGCACTTTTGTGCTCCCTCTCTGCTTTGGCTCCATCAACGGTGGTGTG 1005

QY 1234 TTTGCTGCTCCAGGTTATCTATGTTGCGTCTCGAGAGGTCACCTTCCAGAAATCCTC 1293

Db 1006 TTTGCTGCTCCAGGTTATCTATGTTGCGTCTCGAGAGGTCACCTTCCAGAAATCCTC 1065

QY 1294 TCCATGATTCATGTCGGAAGCAGCACTCTCTACCAAGCTGTATGTTTTCGACCCCTTG 1353

Db 1066 TCCATGATTCATGTCGGAAGCAGCACTCTCTACCAAGCTGTATGTTTTCGACCCCTTG 1125

QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCTCCTCAGTTTGGC 1413

Db 1126 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCTCCTCAGTTTGGC 1185

QY 1414 AGGTGGCTTTTATTTGGGCTGGCAGTCTCTGGCTGATTTATCTTCGATACAAATGCCA 1473

Db 1186 AGGTGGCTTTTATTTGGGCTGGCAGTCTCTGGCTGATTTATCTTCGATACAAATGCCA 1245

QY 1474 GATATGATCGTCCCTTCAAGTGCGCACTGTTCACTCCAGCTTTGTTTCTTCCATGTC 1533

Db 1246 GATATGATCGTCCCTTCAAGTGCGCACTGTTCACTCCAGCTTTGTTTCTTCCATGTC 1305

QY 1534 CTCCTTCATGGTGGCTTTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1593

Db 1306 CTCCTTCATGGTGGCTTTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1365

QY 1594 ATCACTCTGACTGAGTCCCTGGCTATTTATCTCTTTTATATATGAGCAAGAACCCAGG 1653

Db 1366 ATCACTCTGACTGAGTCCCTGGCTATTTATCTCTTTTATATATGAGCAAGAACCCAGG 1425

QY 1654 TGGTTTGAATAATGTCAGAGAAATTAACAGAAATTAACAGAAATTAACAGAAATTAACAG 1713

Db 1426 TGGTTTGAATAATGTCAGAGAAATTAACAGAAATTAACAGAAATTAACAGAAATTAACAG 1485

QY 1714 CCAGAAGAGATAAGTTATGAACCTTAATGGACTTTGAGATCTTGGCAATCTGCCAAGGGA 1773

Db 1486 CCAGAAGAGATAAGTTATGAACCTTAATGGACTTTGAGATCTTGGCAATCTGCCAAGGGA 1545

QY 1774 GACAAATAGGATTTTACTTCAATTTCTGAAAGCTTAGAGAAATTAACAACTTTGGTG 1833

Db 1546 GACAAATAGGATTTTACTTCAATTTCTGAAAGCTTAGAGAAATTAACAACTTTGGTG 1605

QY 1834 ATAAACAAAGAGTCAAGTTATTTTATTCATATATTTTATGACATATTCGAACTAAATTTCT 1893

Db 1606 ATAAACAAAGAGTCAAGTTATTTTATTCATATATTTTATGACATATTCGAACTAAATTTCT 1665

QY 1894 AAGAAATTTAGTTATATCTATGTTATAGAAAGTGAATATGCAAGTTATTTCTATGAG 1953

Db 1666 AAGAAATTTAGTTATATCTATGTTATAGAAAGTGAATATGCAAGTTATTTCTATGAG 1725

QY 1954 TCGCACAATCTTGAGTCTCTGATACCTTCTATGAGGTTAGAGAAAGACTAGACAA 2013

Db 1726 TCGCACAATCTTGAGTCTCTGATACCTTCTATGAGGTTAGAGAAAGACTAGACAA 1785

QY 2014 TTACTATGTTGATCTCTACAACTATGTTAGCAGGCAAGAACTTTCAATTTGAAG 2073

Db 1786 TTACTATGTTGATCTCTACAACTATGTTAGCAGGCAAGAACTTTCAATTTGAAG 1845

QY 2074 ACTGAGATTTTCTGATATATGTTTCTGAAAGTGGTTTATGAGTGGTTTACACACTACAGATGTCT 2133

Db 1846 ACTGAGATTTTCTGATATATGTTTCTGAAAGTGGTTTATGAGTGGTTTACACACTACAGATGTCT 1905

QY 2134 ATACTGTGAAAGTGGTTTCTCAATTTCTGAAAAAGGATACATCATGATTTATGCAAGAG 2193

Db 1906 ATACTGTGAAAGTGGTTTCTCAATTTCTGAAAAAGGATACATCATGATTTATGCAAGAG 1965

QY 2194 GAGAGAA 2200

Db 1966 GAGAGAA 1972

RESULT 6

US-10-163-866-29

; Sequence 29, Application US/10163866

; Publication No. US20030027188A1

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX02-080C

; CURRENT APPLICATION NUMBER: US/10/163,866

; CURRENT FILING DATE: 2002-06-05

; PRIOR APPLICATION NUMBER: US 60/296,076

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/328,605

; PRIOR FILING DATE: 2001-10-10

; PRIOR APPLICATION NUMBER: US 60/338,733

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/357,253

; PRIOR FILING DATE: 2002-02-15

; PRIOR APPLICATION NUMBER: US 60/357,600

; PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 29

; LENGTH: 2482

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-163-866-29

Query Match 76.9%; Score 1722; DB 9; Length 2482;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1772; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393

Db 337 GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 396

QY 394 ACCATCATTTGGAGCAGCAATCTTCATCTCTCTTAAGGGGTGCTCCAGAACACGGCAGC 453

Db 397 ACCATCATTTGGAGCAGCAATCTTCATCTCTCTTAAGGGGTGCTCCAGAACACGGCAGC 456

QY 454 GTGGGATGCTCTGACCATCTGGAGCGTGTGTGGGTCTCTGTCACTATTTGGAGCTTTG 513

Db 457 GTGGGATGCTCTGACCATCTGGAGCGTGTGTGGGTCTCTGTCACTATTTGGAGCTTTG 516

QY 514 TCTTATGCTGAATTTGGGACACACTATAAGAAATCTGGAGGTCATTACACATATATTTG 573

Db 517 TCTTATGCTGAATTTGGGACACACTATAAGAAATCTGGAGGTCATTACACATATATTTG 576

QY 574 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGC 633

Db 577 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGC 636

QY 634 CTGCAAGTCTGCTGTGTGATATCCCTGGCAATTTGGACGCTACATTTGGAACCATTTTT 693

Db 637 CTGCAAGTCTGCTGTGTGATATCCCTGGCAATTTGGACGCTACATTTGGAACCATTTTT 696

QY 694 ATTCAATGTAATCCCTGAACTTGGATCAAGCTCATTTACAGCTGTGGGCAATACTGTA 753

Db 697 ATTCAATGTAATCCCTGAACTTGGATCAAGCTCATTTACAGCTGTGGGCAATACTGTA 756

QY 754 GTGATGGTCTTAATAGCATGATGTCTAGCTGGAGCGCCGGATCCAGATTTTCTTAACC 813

Db 757 GTGATGGTCTTAATAGCATGATGTCTAGCTGGAGCGCCGGATCCAGATTTTCTTAACC 816

QY 814 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGACGATAATTTAA 873

Db 817 TTTTGAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGACGATAATTTAA 876

QY 874 GGTCAAAACGCAAGCTTTTAAAGACGCCCTTTTCAGGAAGAGATTTCAAGTATTACGGGTTG 933

Db 877 GGTCAAAACGCAAGCTTTTAAAGACGCCCTTTTCAGGAAGAGATTTCAAGTATTACGGGTTG 936

Db 1081 TCATTAGCAGTCCGATCTTTGTCCTCTCTCTGCTTTGGCTCCATGAACGGTGGTG 1140
Qy 1234 TTTGCTGCTCCAGGTATTTATGTTGGCTCTCGAGAGGTCACCTTCAGAAATCCCTC 1293
Db 1141 TTTGCTGCTCCAGGTATTTATGTTGGCTCTCGAGAGGTCACCTTCAGAAATCCCTC 1200
Qy 1294 TCCATGATTCATGTCGCGAAGCAGACACCTCTACAGCTGTTATGTTTGGCACCCCTTG 1353
Db 1201 TCCATGATTCATGTCGCGAAGCAGACACCTCTACAGCTGTTATGTTTGGCACCCCTTG 1260
Qy 1354 ACAATGATAAGTCTCTCTCGAGAGCTCGACAGTCTTTTGAATTTCCCTCAAGTTTGGC 1413
Db 1261 ACAATGATAAGTCTCTCTCGAGAGCTCGACAGTCTTTTGAATTTCCCTCAAGTTTGGC 1320
Qy 1414 AGGTGGCTTTTATTTGGGCTGGCAGTGTCTGGGCTGATTTATCTTCGATACAAATGCCA 1473
Db 1321 AGGTGGCTTTTATTTGGGCTGGCAGTGTCTGGGCTGATTTATCTTCGATACAAATGCCA 1380
Qy 1474 GATATGCATCTCTCTCAAGGTGCCACTGTTTCAATCCAGCTTTGTTTCCCTTCACATGC 1533
Db 1381 GATATGCATCTCTCTCAAGGTGCCACTGTTTCAATCCAGCTTTGTTTCCCTTCACATGC 1440
Qy 1534 CTCTTCATGTTGGCTTTCCCTCTATTCGACCCCATTTAGTACAGGATTTGGCTTCGTC 1593
Db 1441 CTCTTCATGTTGGCTTTCCCTCTATTCGACCCCATTTAGTACAGGATTTGGCTTCGTC 1500
Qy 1594 ATCACTCTGACTGGAGTCCCTGGGTATTAATCTCTTTATTTATGGACAAGAACCCAGG 1653
Db 1501 ATCACTCTGACTGGAGTCCCTGGGTATTAATCTCTTTATTTATGGACAAGAACCCAGG 1560
Qy 1654 TGGTTAGATAATGTCAG 1672
Db 1561 TGGTTAGATAATGTCAG 1579

RESULT 10

us-10-163-866-52
; Sequence 52, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; FILE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; TITLE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-52

Query Match 58.0%; Score 1299; DB 9; Length 1528;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 334 GAGAAAGTCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATATCATTTGGC 393
Db 128 GAGAAAGTCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATATCATTTGGC 187
Qy 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCAAGGCGTCTCCAGAACACGGGAGC 453
Db 1268 GATATGCATCGCTCTTTCAAGGTGGAGTGTTCATCCAGCTTTGTTTCTTCACATGC 1327

Db 188 ACCATCATTTGGAGCAGGAATCTTTCATCTCTCTAAGGCGTGTCTCCAGAAACACGGCAGC 247
Qy 454 GTGGCAGTCTCTCGACCATCTGACCGTGTGTGGGTCTGTCTCACTATTTTGGAGCTTTG 513
Db 248 GTGGCAGTCTCTCGACCATCTGACCGTGTGTGGGTCTGTCTCACTATTTTGGAGCTTTG 307
Qy 514 TCTTATGCTGAATTTGGGAACAACCTATAAAGAAATCTGGAGGTCAATTACACATATATTTG 573
Db 308 TCTTATGCTGAATTTGGGAACAACCTATAAAGAAATCTGGAGGTCAATTACACATATATTTG 367
Qy 574 GAAGCTTTTGGTCCATACCGAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATACGC 633
Db 368 GAAGCTTTTGGTCCATACCGAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATACGC 427
Qy 634 CCTCAGCTACTGCTGTGATATCCCTGGCATTTTGGAGCGTACATTTCTGGAACCATTTT 693
Db 428 CCTCAGCTACTGCTGTGATATCCCTGGCATTTTGGAGCGTACATTTCTGGAACCATTTT 487
Qy 694 ATTCAATGTGAATCCCTGAACTTGGATCGGATCAAGCTCATTTACAGCTGTGGGCATACGT 753
Db 488 ATTCAATGTGAATCCCTGAACTTGGATCGGATCAAGCTCATTTACAGCTGTGGGCATACGT 547
Qy 754 GTGATGCTCTTAATAGCATGAGTGTGAGCGTGTGAGCGTGTGAGCGTGTGAGCGTGTG 813
Db 548 GTGATGCTCTTAATAGCATGAGTGTGAGCGTGTGAGCGTGTGAGCGTGTGAGCGTGTG 607
Qy 814 TTTTGAAGCTCAGCAATTTCTGATAATATATAGTCCCTGGAGTTATGAGCTTAATTA 873
Db 608 TTTTGAAGCTCAGCAATTTCTGATAATATATAGTCCCTGGAGTTATGAGCTTAATTA 667
Qy 874 GGTCAAAACGAGAACCTTTAAAGACGCTTTTAAAGACGCTTTTAAAGACGCTTTTAAAG 933
Db 668 GGTCAAAACGAGAACCTTTAAAGACGCTTTTAAAGACGCTTTTAAAGACGCTTTTAAAG 727
Qy 934 CCATGCGCTTTTATGGAATGTATGATATGCTGCTGCTTTTACCTCAACTTTGTT 993
Db 728 CCATGCGCTTTTATGGAATGTATGATATGCTGCTGCTTTTACCTCAACTTTGTT 787
Qy 994 ACTGAAGAGTGAAGAACCTTGAAGAACCTTCCCTTGCATATATGATATATATATATATAT 1053
Db 788 ACTGAAGAGTGAAGAACCTTGAAGAACCTTCCCTTGCATATATGATATATATATATATAT 847
Qy 1054 ATTGTCAACCATTTGCTGTGACAAATGTGGCTACTTTACGACCAATTAATCTCAG 1113
Db 848 ATTGTCAACCATTTGCTGTGACAAATGTGGCTACTTTACGACCAATTAATCTCAG 907
Qy 1114 GAGCTGCTGCTTCAATGCAATGAGTGGAGTGACCTTTTCTGAGCGCTACTGGGAAATTC 1173
Db 908 GAGCTGCTGCTTCAATGCAATGAGTGGAGTGACCTTTTCTGAGCGCTACTGGGAAATTC 967
Qy 1174 TCATTAGCAGTTCGATCTTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1233
Db 968 TCATTAGCAGTTCGATCTTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1027
Qy 1234 TTTGCTGCTCCAGGTTATTTCTATGTTGGTCTCGAGAGGTCACCTTTCAGAAATCCCT 1293
Db 1028 TTTGCTGCTCCAGGTTATTTCTATGTTGGTCTCGAGAGGTCACCTTTCAGAAATCCCT 1087
Qy 1294 TCCATGATTCATGTCGCAAGCAGACACTCCTCTACAGCTGTTATTTGTTTGGACCTTTG 1353
Db 1088 TCCATGATTCATGTCGCAAGCAGACACTCCTCTACAGCTGTTATTTGTTTGGACCTTTG 1147
Qy 1354 ACAATGATAATGCTCTCTCTCGAGACCTCGACAGCTTTTGAATTTTCTCAGTTTGGC 1413
Db 1148 ACAATGATAATGCTCTCTCTCGAGACCTCGACAGCTTTTGAATTTTCTCAGTTTGGC 1207
Qy 1414 AGGTGGCTTTTATTTGGCTGGCAGTGTGCTGGGTGATTTATCTTCGATACAAATGCCA 1473
Db 1208 AGGTGGCTTTTATTTGGCTGGCAGTGTGCTGGGTGATTTATCTTCGATACAAATGCCA 1267
Qy 1474 GATATGCATCGCTCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCACATGC 1533
Db 1268 GATATGCATCGCTCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCACATGC 1327

QY 1534 CTCTTCATGTTGCCCTTCCCTCTAFTTCGGACCCCAATTTAGTACAGGGATGGCTTCGTC 1593
Db 1328 CTCTTCATGTTGCCCTTCCCTCTAFTTCGGACCCCAATTTAGTACAGGGATGGCTTCGTC 1387
QY 1594 ATCACTCTGACTGGAGTCCCTCGGTATTAATCTCTTTATATATGGACAGAAACCCAGG 1653
Db 1388 ATCACTCTGACTGGAGTCCCTCGGTATTAATCTCTTTATATATGGACAGAAACCCAGG 1447
QY 1654 TGGTTTGAATAATGTCCAGAGAAATAAACAGAACATTTACAAATAATCTGGAAGTTGA 1713
Db 1448 TGGTTTGAATAATGTCCAGAGAAATAAACAGAACATTTACAAATAATCTGGAAGTTGA 1507
QY 1714 CCAGAAGAAGATAAGTTATGA 1734
Db 1508 CCAGAAGAAGATAAGTTATGA 1528

RESULT 11
US-10-163-866-32
; Sequence 53, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLIC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 1268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-53

Query Match 38.8%; Score 868; DB 9; Length 1268;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 868; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Db 128 GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 187
QY 394 ACCATCTTGGAGAGGAATCTTCATCTCTCTTAAGGCGTGTCTCCAGAACACGGCAGC 453
Db 188 ACCATCTTGGAGAGGAATCTTCATCTCTCTTAAGGCGTGTCTCCAGAACACGGCAGC 247
QY 454 GTGGGCATCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGCTACTATTTGGAGCTTTG 513
Db 248 GTGGGCATCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGCTACTATTTGGAGCTTTG 307
QY 514 TCTTATGCTGAATTTGGGAACAACACTATAAGAAATCTGAGGTCATTACACATATATTTTG 573
Db 308 TCTTATGCTGAATTTGGGAACAACACTATAAGAAATCTGAGGTCATTACACATATATTTTG 367
QY 574 GAAGTCTTTGGTCCATTACACCTTTTGTACAGTCTGGGTGGAAGTCTCTCAATATACGC 633
Db 368 GAAGTCTTTGGTCCATTACACCTTTTGTACAGTCTGGGTGGAAGTCTCTCAATATACGC 427
QY 634 CTGTCAGCTACTGCTGTATATCCCTGGCAATTTGGACGCTACATTTCTGGAACCAATTTT 693
Db 428 CCTGCAGCTACTGCTGTATATCCCTGGCAATTTGGACGCTACATTTCTGGAACCAATTTT 487

QY 694 ATTCAATGTGAAATCCCTGAACTTGGGATCAAGCTCAATTTACAGCTGTGGGCATAACTGTA 753
Db 488 ATTCAATGTGAAATCCCTGAACTTGGGATCAAGCTCAATTTACAGCTGTGGGCATAACTGTA 547
QY 754 GTGATGGTCTTAAATAGCATGAGTGTCAAGTGGAGCGGCCGATCCAGATTTTCTTAACC 813
Db 548 GTGATGGTCTTAAATAGCATGAGTGTCAAGTGGAGCGGCCGATCCAGATTTTCTTAACC 607
QY 814 TTTTCCAAGCTCACAGCAATTTTAAAGACGCCCTTTTCAGGAAGAGATTTCAAGTATTTACCGGTTG 873
Db 608 TTTTCCAAGCTCACAGCAATTTTAAAGACGCCCTTTTCAGGAAGAGATTTCAAGTATTTACCGGTTG 667
QY 874 GGTCAAAACGAGCAATTTTAAAGACGCCCTTTTCAGGAAGAGATTTCAAGTATTTACCGGTTG 933
Db 668 GGTCAAAACGAGCAATTTTAAAGACGCCCTTTTCAGGAAGAGATTTCAAGTATTTACCGGTTG 727
QY 934 CCACTGGCTTTTATATGGAATGTATGATATGCTGCTGGTGTGCTTACCTCAACTTTGTT 993
Db 728 CCACTGGCTTTTATATGGAATGTATGATATGCTGCTGGTGTGCTTACCTCAACTTTGTT 787
QY 994 ACTGAAGAGTAGAAAACCCCTGAAAAACCACTTCCCTTTGCAATATGATATCCATGGCC 1053
Db 788 ACTGAAGAGTAGAAAACCCCTGAAAAACCACTTCCCTTTGCAATATGATATCCATGGCC 847
QY 1054 ATTGCACCAATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCACTTAATGCTGAG 1113
Db 848 ATTGCACCAATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCACTTAATGCTGAG 907
QY 1114 GAGCTGCTCTTTCAAAATGCAATGCACTGGCAGTACTTTCTGAGCGGCTACTGGGAAATTTTC 1173
Db 908 GAGCTGCTCTTTCAAAATGCAATGCACTGGCAGTACTTTCTGAGCGGCTACTGGGAAATTTTC 967
QY 1174 TCATTAGCAGTTCGCGATCTTTTGTGGCC 1201
Db 968 TCATTAGCAGTTCGCGATCTTTTGTGGCC 995

RESULT 12
US-10-163-866-32
; Sequence 32, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLIC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-32

Query Match 18.5%; Score 415; DB 9; Length 520;
Best Local Similarity 100.0%; Pred. No. 1.2e-201;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Db 106 GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 165

QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTAAGGCGTGTCTCCAGAACACGGCAGC 453
Db 166 ACCATCATTTGGAGCAGGAATCTTCATCTCTAAGGCGTGTCTCCAGAACACGGCAGC 225
QY 454 GTGGGCATGCTCTGACCATCTGACGGGTGTGGGGTCTGTCACTATTGGAGCTTTG 513
Db 226 GTGGGCATGCTCTGACCATCTGACGGGTGTGGGGTCTGTCACTATTGGAGCTTTG 285
QY 514 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGTCACTACATATATTTG 573
Db 286 TCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGTCACTACATATATTTG 345
QY 574 GAAGTCTTTGGTCCATTACCAAGCTTTTGTACGAGTCTGGTGGAACTCTCTCAATAACGC 633
Db 346 GAAGTCTTTGGTCCATTACCAAGCTTTTGTACGAGTCTGGTGGAACTCTCTCAATAACGC 405
QY 634 CCTGCAGTACTGCTGTGATATCCCTGGCATTTGGAGCTACATCTGGAACATTTT 693
Db 406 CCTGCAGTACTGCTGTGATATCCCTGGCATTTGGAGCTACATCTGGAACATTTT 465
QY 694 ATCAATGTGAATCCCTGAACCTTGGCATCAAGCTCATTACAGCTGTGGGCATAA 748
Db 466 ATCAATGTGAATCCCTGAACCTTGGCATCAAGCTCATTACAGCTGTGGGCATAA 520

RESULT 13

US-09-854-133-442
; Sequence 442, Application US/09854133
; Patent No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Henderson, Raodoh
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-133-442

Query Match 12.6%; Score 281; DB 9; Length 337;
Best Local Similarity 100.0%; Pred. No. 3.8e-133;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAGTTGAAGTGAGCAGAGATCATGCCAGCTGGTGACAGTGTCTGTCTCAAC 60
Db 4 GGAGTTGAAGTGAGCAGAGATCATGCCAGCTGGTGACAGTGTCTGTCTCAAC 63
QY 61 AGAATTAAGAAAAAGAAAAAGAGAGAGAGAAATCCAGGCCAATTGTGG 120
Db 64 AGAATTAAGAAAAAGAAAAAGAGAGAGAGAAATCCAGGCCAATTGTGG 123
QY 121 CATAGATTTTATCATATCTGGATTTTGGATTTTGTCTCTCATCTGATTC 180
Db 124 CATAGATTTTATCATATCTGGATTTTGGATTTTGTCTCTCATCTGATTC 183
QY 181 GGAAGCCCTTCCCTGGGCAACAGAGAGAGAAATCCAGGCCAATTGTGG 240
Db 184 GGAAGCCCTTGGTGTCACCATCTCCAAAGAGAGGTACCTGCGAGGAAATGTTAACGGG 243
QY 241 AGGCTCCCTTCCCTGGGCAACAGAGAGAGAAATCCAGGCCAATTGTGG 281
Db 244 AGGCTCCCTTCCCTGGGCAACAGAGAGAGAAATGTTAACGGG 284

RESULT 14

US-09-738-973-442
; Sequence 442, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-973-442

Query Match 12.6%; Score 281; DB 10; Length 337;
Best Local Similarity 100.0%; Pred. No. 3.8e-133;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAGTTGAAGTGAGCAGAGATCATGCCAGCTGGTGACAGTGTCTGTCTCAAC 60
Db 4 GGAGTTGAAGTGAGCAGAGATCATGCCAGCTGGTGACAGTGTCTGTCTCAAC 63
QY 61 AGAATTAAGAAAAAGAAAAAGAGAGAGAGAAATCCAGGCCAATTGTGG 120
Db 64 AGAATTAAGAAAAAGAAAAAGAGAGAGAGAAATCCAGGCCAATTGTGG 123
QY 121 CATAGATTTTATCATATCTGGATTTTGGATTTTGTCTCTCATCTGATTC 180
Db 124 CATAGATTTTATCATATCTGGATTTTGGATTTTGTCTCTCATCTGATTC 183
QY 181 GGAAGCCCTTCCCTGGGCAACAGAGAGAGAAATCCAGGCCAATTGTGG 240
Db 184 GGAAGCCCTTGGTGTCACCATCTCCAAAGAGAGGTACCTGCGAGGAAATGTTAACGGG 243
QY 241 AGGCTCCCTTCCCTGGGCAACAGAGAGAGAAATCCAGGCCAATTGTGG 281
Db 244 AGGCTCCCTTCCCTGGGCAACAGAGAGAGAAATGTTAACGGG 284

RESULT 15

US-10-046-935-1307
; Sequence 1307, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aljun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1307
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9, 19, 461, 497, 500, 502
; OTHER INFORMATION: n = A,T,C or G
US-10-046-935-1307

Query Match 4.8%; Score 107; DB 9; Length 572;
Best Local Similarity 100.0%; Pred. No. 3.3e-44;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GAGAAAGTCAGCTCAAGAGGAAAGTCACCTTACTGAGGGAGTCTCCATTATCATTTGGC 393
DB 354 GAGAAAGTCAGCTCAAGAGGAAAGTCACCTTACTGAGGGAGTCTCCATTATCATTTGGC 413
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGGCGTGCTCCA 440
DB 414 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGGCGTGCTCCA 460

Search completed: April 16, 2003, 16:06:46
Job time : 236 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - protein search, using frame_plus_n2p model
Run on: April 16, 2003, 14:10:10 ; Search time 71 Seconds
(without alignments)
8404.173 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 4120
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-O=Ccp2_1/USPTO.spool/US09667170/runat_04042003_090917_20278/app_query.fasta_1.2375
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09667170.ecgn_1.1.79 -runat_04042003_090917_20278 -NCP=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB ID	Description
1	61.1	2519	521	22	Human cysteine/Glu
2	52.6	2168	424	22	Human EST encoded
3	29.4	1210.5	507	21	Human L-type amino
4	29.2	1204	512	21	Rat L-type amino a
5	29.2	1204	512	21	Rat neutral amino
6	27.7	1140	507	22	Human 4P2 light ch
7	27.5	1131.5	511	20	Human amino acid p
8	27.5	1131.5	511	20	Human amino acid p
9	27.5	1131.5	511	20	Human momp-2 prote
10	27.0	1113.5	505	22	Drosophila melanog
11	26.9	1108	533	22	Rat L-type amino a
12	26.8	1106	499	22	Drosophila melanog
13	26.7	1101.5	535	21	Human ORFX ORF2900
14	26.7	1101.5	535	21	Human membrane tra
15	26.7	1101.5	535	21	Human L-type amino
16	26.5	1091	517	22	Drosophila melanog
17	26.3	1083	564	22	Human amino acid t
18	26.1	1075.5	500	22	Human hydrophobic
19	26.1	1075.5	487	22	Amino acid sequenc
20	25.5	1051.5	530	22	Drosophila melanog
21	25.2	1039	523	22	Human secreted pro
22	22.8	938.5	445	22	Human secreted pro
23	18.0	741.5	352	21	Human pancreatic c
24	17.8	732	414	21	Human secreted pro
25	17.7	730	370	21	Human secreted pro
26	17.4	714	386	22	Novel human diago
27	14.6	603	377	21	Human EL6H protein
28	13.5	556.5	389	22	Human hydrophobic
29	12.7	524.5	245	19	Human tumour-assoc
30	11.5	475	194	23	Human polypeptide
31	10.1	415.5	222	23	Human polypeptide
32	9.9	406.5	179	22	Peptide #5187 enco
33	9.9	406.5	179	22	Human brain expres
34	9.9	406.5	179	22	Human bone marrow
35	9.9	406.5	179	22	Peptide #5058 enco
36	9.9	406.5	179	22	Peptide #5121 enco
37	9.9	406.5	179	23	Human peptid enco
38	8.0	331	463	23	Listeria monocytog
39	7.8	323	462	22	E. coli growth and
40	7.8	322	456	22	Enterococcus faeca
41	7.8	322	463	22	Enterococcus faeca
42	7.6	312.5	486	23	Streptococcus poly
43	7.5	307.5	583	21	Arabidopsis thalia
44	7.5	307.5	595	21	Arabidopsis thalia
45	7.4	305.5	538	21	Arabidopsis thalia

ALIGNMENTS

RESULT 1
ABBI1957
ID ABBI1957 standard; peptide; 521 AA.
XX AC ABBI1957;
XX AC
XX 11-JAN-2002 (first entry)
XX Human cysteine/Glu transporter homologue, SEQ ID NO:2327.
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibit; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;

KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnerary; antiulcer.

OS Homo sapiens.

XX W0200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US03800.

XX 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEO INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

DR N-PSDB; ABA09201.

XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer.

XX Claim 20; Page 286; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.

XX Sequence 521 AA;

XX Alignment Scores:

Pred. No.: 3,366-276 Length: 521

Score: 2519.00 Matches: 498

Percent Similarity: 96.51% Conservative: 0

Best Local Similarity: 96.51% Mismatches: 0
 Query Match: 61.14% Indels: 18
 DB: 22 Caps: 1

US-09-667-170A-440 (1-2239) x ABB11957 (1-521)

QY 184 AAGCTGTGTGTCCACCATCTCCAAAGGAGGTTACCTGCGAGGAAATGTTAAGGGAGG 243

Db 24 LysProValValSerThrIleSerLysGlyTyrLeuGlnGlyAsnValAsnGlyArg 43

QY 244 CTGCTTCCCTGGGCAACAGGAGCCACCTGGCGAGGACGCCCTTTTCAGGAAGAGAGGCC 303

Db 44 LeuProSerLeuGlyAsnLysGluProProGlyGln----- 55

QY 304 TTTTCAGGACAGACAGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTCACT 363

Db 56 -----GluLysValGlnLeuLysArgLysValThr 65

QY 364 TTACTGAGGGAGTCTCCATATATATGCGACCATCATTTGGAGGAGGAAATCTTCATCTCT 423

Db 66 LeuLeuArgGlyValSerIleIleGlyThrIleIleGlyAlaGlyIlePheIleSer 85

QY 424 CCTAAGGCGTGCCTCCAGAACACAGCGGTGGCATGTCTCTGACCATCTGGACGGTG 483

Db 86 ProLysGlyValLeuGlnAsnThrGlySerValGlyMetSerLeuThrIleIlePheVal 105

QY 484 TCTGGGCTCTGTCTACTATTTGGAGCTTTGTCTTTATCTGCTGAATTTGGAAACAATATAAG 543

Db 106 CysGlyValLeuSerLeuPheGlyAlaLeuSerTyrAlaGluLeuGlyThrIleLys 125

QY 544 AAATCTGGAGTCTATTCACATATATTTTGGAGTCTTTGGTCCATTTACAGCTTTTGTGA 603

Db 126 LysSerGlyGlyHisTyrThrTyrIleLeuGluValPheGlyProLeuProAlaPheVal 145

QY 604 CGAGTCTGGGTGGAACTCCTCATATACGCCCTGACCTACTGCTGTGATATCCTCGCA 663

Db 146 ArgValTrpValGluLeuLeuIleIleArgProAlaIleAlaValIleSerLeuAla 165

QY 664 TTTGGAGCTACATCTCTGGAACCAATTTTATTCATTTGTAATTCCTGAACTTGGCATC 723

Db 166 PheGlyArgTyrIleLeuGluProPhePheIleGlnCysGluIleProGluLeuAlaIle 185

QY 724 AAGCTCATTACAGCTGTGGGCATACCTGATGCTGTCTTAAATAGCATGAGTGTCAGC 783

Db 186 LysLeuIleThrAlaValGlyIleThrValValMetValLeuAsnSerMetSerValSer 205

QY 784 TGGAGCGCGCGGATCCAGATTTTCTTAACTTTTGAAGCTTCACAGCAATTTGATAATT 843

Db 206 TrpSerAlaArgIleGlnIlePheLeuThrPheCysLysLeuThrAlaIleLeuIle 225

QY 844 ATAGTCCCTGGAGTATGACAGCTAATTAAGGTCAAAGCGAGAGACTTTAAGAGCGCTTT 903

Db 226 IleValProGlyValMetGlnLeuIleLysGlyGlnThrGlnAsnPheLysAspAlaPhe 245

QY 904 TCAGGAAGAGATTCAGATTTACCGGTGTGCACTGGCTTTTATTATGAAATGATGCA 963

Db 246 SerGlyArgAspSerSerIleThrArgLeuProLeuAlaPheTyrTyrGlyMetTyrAla 265

QY 964 TATGCTGGCTGGTCTTACCTCAACTTTGTTACTGAGAAGTAGAAAACCTGAAAAAAC 1023

Db 266 TyrAlaGlyTrpPheTyrLeuAsnPheValThrGluGluValGluAsnProGluLysThr 285

QY 1024 ATTCCTCTTGAATATGATATATCATGCGCATTTGTCACCATTTGGCTATGCTGACAAAT 1083

Db 286 IleProLeuAlaIleCysIleSerMetAlaIleValThrIleGlyTyrValLeuThrAsn 305

QY 1084 GTGGCTACTTTACGACCATTAATGCTGAGAGAGCTGCTGCTTTCAAATGAGTGGCAGTG 1143

Db 306 ValAlaTyrPheThrThrIleAsnAlaGluLeuLeuLeuSerAsnAlaValAlaVal 325

QY 1144 ACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTACAGTTCGGATCTTTGTTGGCTC 1203

Db 326 ThrPheSerGluArgLeuLeuGlyAsnPheSerLeuAlaValProIlePheValAlaLeu 345

QY 1204 TCCTGCTTTGGCTCCATGAACGGTGTGTCTGCTGCTCCAGGTTATCTATGTCG 1263
 DB 346 SerCysPheGlySerMetAsnGlyValPheAlaValSerArgLeuPheTyrValAla 365
 QY 1264 TCTCGAGGGTCACTTCCTCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACTCCT 1323
 DB 366 SerArgGluGlyHisLeuProGluLeuSerMetIleHisValArgLysHisThrPro 385
 QY 1324 CTACACAGCTGTATGTTTGGACCCCTTCACATGATGATGCTCTCTCTGAGACCTC 1383
 DB 386 LeuProAlaValIleValLeuHisProLeuThrMetIleMetLeuPheSerGlyAspLeu 405
 QY 1384 GACAGCTTTTGAATTTCTCAGTTTTCAGGTGGCTTTTATTTGGCTGGCAGTTGCT 1443
 DB 406 AspSerLeuLeuAsnPheLeuSerPheAlaArgTrpLeuPheIleGlyLeuAlaValAla 425
 QY 1444 GGCTGATTTATCTTCATGATACAAATGCCAGATATGATGCTGCTCTTCAAGGTGCCACTG 1503
 DB 426 GlyLeuIleTyrLeuArgTyrLysCysProAspMetHisArgProPheLysValProLeu 445
 QY 1504 TTATCCACAGCTTTGTTTCTCTTCACATGCTCTTCTATGTTGGCTTCCCTCTATTGG 1563
 DB 446 PheIleProAlaLeuPheSerPheThrCysLeuPheMetValAlaLeuSerLeuTyrSer 465
 QY 1564 GACCATTTAGTACAGGATTTGCTTCTGTCATCCTGCTGCTGCTGCTGCTGCTGCTGCT 1623
 DB 466 AspProPheSerThrGlyIleGlyPheValIleThrLeuThrGlyValProAlaTyrTyr 485
 QY 1624 CTCTTTATTTATATGGACAGAACCCAGGTGGTTTGTAGATATATGTCAGAGAAATAACC 1683
 DB 486 LeuPheIleIleThrAspLysLysProArgTrpPheArgIleMetSerGluLysIleThr 505
 QY 1684 AGACATTTACAAATAATCTGGAAGTTGTACCAGAGAGAGATAGCTTA 1731
 DB 506 ArgThrLeuGlnIleIleLeuGluValProGluGluAspLysLeu 521

RESULT 2
 AAM23914
 ID AAM23914 standard; Protein; 424 AA.
 AC AAM23914;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST encoded protein SEQ ID NO: 1439.
 XX
 KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition.
 XX
 OS Homo sapiens.
 XX
 PN WO200154477-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US02687.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 XX
 DR WPI: 2001-476164/51.
 DR N-PSDB: AAH98573.
 XX

PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -
 XX
 PS Claim 20; Page 1001-1002; 1275pp; English.
 XX

The present invention provides the protein and coding sequences of novel
 proteins from a variety of organisms, including human, dog, cat, horse,
 cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 urchin and tomato. These were derived from expressed sequence tags (ESTs)
 from the organism of interest. They can be used in diagnostics,
 for forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.

SQ Sequence 424 AA;

Alignment Scores:

Pred. No.: 2, 08e-236 Length: 424
 Score: 2168.00 Matches: 424
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 52.62% Indels: 0
 DB: Gaps: 0

US-09-667-170A-440 (1-2239) x AAM23914 (1-424)

QY 460 ATGCTCTGACATCTGACGCTGTGGGTCTCTGCTACATTTTGGAGCTTTGCTTAT 519
 DB 1 MetSerLeuThrIleTrpThrValCysGlyValLeuSerLeuPheGlyAlaLeuSerTyr 20
 QY 520 GCTGAATTTGGACAACTATAAGAAATCTCGAGGTCTTACACATATATTTTGAAGTC 579
 DB 21 AlaGluLeuGlyThrIleLysLysSerGlyGlyHisTyrThrTyrIleLeuGluVal 40
 QY 580 TTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAACTCTCAATAACGCCCTGCA 639
 DB 41 PheGlyProLeuProAlaPheValArgValTrpValGluLeuLeuIleIleArgProAla 60
 QY 640 GCTACTGCTGTGATATCCTCGCATTTTGGAGCTTACATTCGTGAACACCTTTTATTCAA 699
 DB 61 AlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProPhePheIleGln 80
 QY 700 TGTGAATCCCTGAACCTTGGCATCAAGCTCATTACAGCTGTGGGCATACTGTAGTATG 759
 DB 81 CysGluIleProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrValValMet 100
 QY 760 GTCCTAAATAGCATGAGTGTACGCTGGAGCCCGCGATCCAGATTTCTTTAACTTTTGC 819
 DB 101 ValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThrPheCys 120
 QY 820 AAGCTCACAGCAATCTGATAATTTATAGTCCCTGGAGTTATGACAGCTAATTAAGGTCAA 879
 DB 121 LysLeuThrAlaIleLeuIleIleValProGlyValMetGlnLeuIleLysGlyGln 140
 QY 880 ACGCAGAACTTTAAAGACGCTTTTTCAGGAAGAGATTCAGATTCACGCGTGGCCACTG 939
 DB 141 ThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerSerIleThrArgLeuProLeu 160
 QY 940 GCTTTTATTATGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTTGTACTGAA 999
 DB 161 AlaPheTyrTyrGlyMetTyrAlaTyrAlaGlyTrpPheTyrLeuAsnPheValThrGlu 180
 QY 1000 GAAGTCAAAACCTGAAAAACCATTCCTTCCATATGATATATCCATGGCCATTGTC 1059
 DB 181 GluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAlaIleVal 200
 QY 1060 ACCATTGGCTATGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTG 1119
 DB 201 ThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrThrIleAsnAlaGluGluLeu 220
 QY 1120 CTGCTTTCAAAATGACAGTGGCAGTACCTTTTCGAGGGCTACTGGGAAATTTCTCATTA 1179
 DB 221 LeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPheSerLeu 240

Db 233 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 252
 QY 961 GCATATGCTGGCTGGTTTACCTCACTTGTACTGAAGAGTAGAAACCCCTGAAAAA 1020
 Db 253 AlaTyrGlyGlyTrpAsnTrpLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
 QY 1021 ACCATTCCCTTGCATATGATATCCATGCGCATATGTCACCATGGCTATGCTGACA 1080
 Db 273 AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValTyrValLeuThr 292
 QY 1081 ATGTGGCTACTTACGACCATTAATGCTGAGGAGTGCCTTTCATATGAGTGGCA 1140
 Db 293 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 312
 QY 1141 GTGACCTTTTTCAGCGCTACTGGGAAATTTCTCATAGAGTTCGATCTTTGTGGC 1200
 Db 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 332
 QY 1201 CTCTCTGCTTGGCTCCATGAAGGCTGGTGTGTGTCTGCTCCAGGTTATTCATGTT 1260
 Db 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPheVal 352
 QY 1261 GGTCTCCAGAGGTCACCTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCACACT 1320
 Db 353 GlySerArgGlyGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372
 QY 1321 CCTCTACACGCTGTTATGTTTTCACCCCTTGACATGATATGCTCTCTCTGAGAC 1380
 Db 373 ProValProSerLeuValPheThrCysValMetThrLeuLeuTyrAlaPheSerLysAsp 392
 QY 1381 CTCGACACTTTTGAATTTCTCAGTTTTCGAGGTGGCTTTTATTTGGCTGGCAGTT 1440
 Db 393 IlePheSerValIleAsnPhePheSerPheAsnTrpLeuCysValAlaLeuAlaIle 412
 QY 1441 GCTGGCTGATTTATCTCCATCAAAATGCCAGATATGATCGTCTCTTCAAGTGGCA 1500
 Db 413 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 432
 QY 1501 CTGTTTCCGACGCTGTTGTTTCTCCATGCTCTCTGATGTTGCTTCCCTCTAT 1560
 Db 433 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheThr 452
 QY 1561 TCGGACCATTTAGTACAGGATGCTGCTCGTATCATCTGCTGAGTCCCTCGTAT 1620
 Db 453 LysThrProValGluCysGlyIleGlyPheThrIleIleLeuSerGlyLeuProValTyr 472
 QY 1621 TATCTCTTTATATATGGGACAAAGAACCCAGGTGTTTAGAATAATGTCAGAGAAATA 1680
 Db 473 PhePheGlyValTrpTrpLysAsnLysProLysTrpLeuLeuGlnGlyIlePheSerThr 492
 QY 1681 ACCAGAACATTAACAATAATCTGGAAGTTGTACCAAGAA 1722
 Db 493 ThrValLeuCysGlnLysLeuMetGlnValValProGlnGlu 506

RESULT 4
 AAY82489
 ID AAY82489 standard; Protein; 512 AA.
 XX
 AC AAY82489;
 XX
 DT 12-JUL-2000 (first entry)
 XX
 DE Rat L-type amino acid transporter 1 protein sequence SEQ ID NO:4.
 XX
 KW L-type amino acid transporter 1; LAT1; 4F2hc; 4F2 heavy chain;
 KW neutral amino acid transporter; tumour; cancer; proliferation;
 KW cell membrane surface 4F2 molecule; anticancer.
 XX
 OS Rattus sp.
 XX
 PN WO200014228-A1.

PD 16-MAR-2000.
 XX
 PF 03-SEP-1999; 99WO-JP04789.
 XX
 PR 03-SEP-1998; 98JP-0249993.
 PR 02-SEP-1999; 99JP-0248546.
 XX
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 PI Endou H, Kanai Y;
 XX
 DR WPI; 2000-256979/22.
 DR N-PSDB; AAA08373.
 XX
 PT Neutral amino acid transporter protein which conjugates with cell
 PT membrane protein 4F2 and operates independently of sodium ions is
 PT useful for screening potential cancer proliferation inhibitors -
 XX
 PS Claim 5; Page 144-147; 189pp; Japanese.
 XX
 CC The present invention describes a cell surface protein which is an
 CC L-type amino acid transporter-1 (LAT1), which mediates the transport
 CC of neutral amino acids, leucine, isoleucine, phenylalanine, methionine,
 CC tyrosine, tryptophan, valine and histidine, into the cell independently
 CC of sodium ions. The LAT1 protein conjugates with the cell membrane
 CC surface molecule 4F2. Antisense DNA or RNA, antibodies, antibody
 CC fragments, and peptide and non-peptide antagonists to LAT1 are useful
 CC as anticancer agents. The present sequence represents rat LAT1, which
 CC is specifically claimed in the present invention.
 SQ Sequence 512 AA;

Alignment Scores:
 Pred. No.: 6,22e-127 Length: 512
 Score: 1204.00 Matches: 229
 Percent Similarity: 67.57% Conservative: 96
 Best Local Similarity: 47.61% Mismatches: 150
 Query Match: 29.22% Indels: 6
 DB: 21 Gaps: 2
 US-09-667-170A-440 (1-2239) x AAY82489 (1-512)

QY 298 GACGCCCTTTTCAGAGAGACGCGCTTTTCAGGAGAGAGAGAGAGAGAGAGAGAGAA 357
 Db 31 GluAlaArgArgGlyAspGlyAlaAspProGluGlyGluGlyValThrLeuGlnArgAsn 50
 QY 358 GTCACCTTACTGAGGGAGTCTCCATTTATTCATTTGGCACCATCATTTGGAGCAGGAATCTTC 417
 Db 51 IleThrLeuIleAsnGlyValAlaIleIleValGlyThrIleIleGlySerGlyIlePhe 70
 QY 418 ATCTCTCTAAGGCGCTGCCAGAACACGGCGCGGCGGCGATGCTCTGACCATCTGG 477
 Db 71 ValThrProThrGlyValLeuLysGluAlaGlySerProGlyLeuSerLeuValValTrp 90
 QY 478 ACGGTGTGTGGGTCTGCTCACTATTGAGCTTTGCTTATGCTGAAATGGGAACAAC 537
 Db 91 AlaValCysGlyValPheSerIleValGlyAlaLeuCysTyrAlaGluLeuGlyThrThr 110
 QY 538 ATAAAGAAATCGGAGGTCAATACACATATATTTGGAGTCTTTGGTCCATTACCACT 597
 Db 111 IleSerLysSerGlyGlyAspTyrAlaTyrMetLeuGluValTyrGlySerLeuProAla 130
 QY 598 TTTGTACGAGTCTGGGTGAACCTCTCAATACCCCTCGAGCTACTGCTGATATCC 657
 Db 131 PheLeuLysLeuTrpIleGluLeuIleIleArgProSerSerGlnTyrIleValAla 150
 QY 658 CTGGCATTTGGAGGTACATTTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCT 717
 Db 151 LeuValPheAlaThrTyrLeuLeuLysProValPheProThrCysProValProGluGlu 170
 QY 718 GCGATCAAGCTCATTTACAGCTGTGGGCATATACTAGTAGTGTGCTCAATAGCATGAGT 777
 Db 171 AlaAlaLysLeuValAlaCysValLeuLeuLeuThrAlaValAsnCysTyrSer 190

Thu Apr 17 09:48:23 2003

us-09-667-170a-440.n2p.rag

QY 778 GTCAGTGGAGCGCCGATCCAGATTTCTTACCTTTTTCAGCTCAGCAATTCG 837
 Db 191 VallysAlaAlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAla 210
 QY 838 ATAATTATAGTCCCTGGAGTTATCGAGTAAATTA-----GGTCAACACGAG 885
 Db 211 LeuileleLeuLeuGlyPheileGlnMetGlyLysAspIleGlyGlnGlyAspAlaSer 230
 QY 886 AACTTT-----AAAGACGCTTTTCAGGAAGAGATTCAAGTATACCGGGTCCCACTG 939
 Db 231 AsnLeuHisGlnLysLeuSerPheGluGlyThrAsnLeuAspValGlyAsnIleValLeu 250
 QY 940 GCTTTTATATGGAATGATATGCTATGCTGCTGGTTTACCTCAACTTTGTTACTGAA 999
 Db 251 AlaLeuTyrrSerGlyLeuPheAlaTyrrGlyGlyTrpAsnTyrrLeuAsnPheValThrGlu 270
 QY 1000 GAAGTAGAACCCTGAACAAACCAATCCCTTCGAATATGATATATCATCGCCATGTC 1059
 Db 271 GluMetIleAsnProtyrArgAsnLeuProLeuAlaIleIleSerLeuProIleVal 290
 QY 1060 ACCATTGGCTATGCTGACAAATGGCTACTTTACGACCAATTAATGCTGAGGAGCTG 1119
 Db 291 ThrLeuValTyrrValLeuThrAsnLeuAlaTyrrPheThrLeuSerThrAsnGlnMet 310
 QY 1120 CTGCTTTCAATGCGAGTGCAGTGCCTTTCTGAGCGGCTACTGGGAAATTTCTCATTA 1179
 Db 311 LeuThrSerGluAlaValAlaValAspPheGlyAsnTyrrHisLeuGlyValMetSerTrp 330
 QY 1180 GCAGTTCCGATCTTGTTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239
 Db 331 IleIleProValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThr 350
 QY 1240 GTCTCCAGGTATTTATGCTGGCTCGAGAGGTCACCTTCACAGAAATCCTCTCCATG 1299
 Db 351 SerSerArgLeuPhePheValGlySerArgGlyHisLeuProSerIleLeuSerMet 370
 QY 1300 ATTCATGTCGCAAGCACATCTCTACAGCTGTTATGTTTTCACCTTTGACAAATG 1359
 Db 371 IleHisProGlnLeuLeuThrProValProSerLeuValPheThrCysValMetThrLeu 390
 QY 1360 ATAATGCTCTCTCTGAGACCTCGACAGCTTTTGAATTTCTCAGTTTTCCTCAGTTT 1419
 Db 391 MetTyrrAlaPheSerArgAspIlePheSerIleAsnPhePheSerPheAsnTrp 410
 QY 1420 CTTTTTATGGCTGCGAGTGTGCTGGCTGATTTATCTTCGATACAAATGCCAGATAG 1479
 Db 411 LeuCysValAlaLeuAlaIleIleGlyMetMetTrpLeuArgPheLysProGluLeu 430
 QY 1480 CATCGCTCTTTCAAGGTGCGACCTGTCATCCAGCTTTGTTTTCCTTCACATGCCCTTC 1539
 Db 431 GluArgProIleLysValAsnLeuAlaLeuProValPhePheIleLeuAlaCysLeuPhe 450
 QY 1540 ATGGTTCCCTTCCCTCTATTCGACCCCAATTTAGTACAGGATTTGGCTTCGTCATCAT 1599
 Db 451 LeuileAlaValSerPheTrpLysThrProLeuGluCysGlyIleGlyPheAlaIleIle 470
 QY 1600 CTGACTGGAGTCCCTGGTATTTATCTCTTTATTTATATGGACAAGAACCCAGGTGGTTT 1659
 Db 471 LeuSerGlyLeuProValTyrrPhePheGlyValTrpTrpLysAsnLysProLysTrpIle 490
 QY 1660 AGAATAATGTCAGAGAAATAACCAAGAACATTAACAATAATCTGGAAGTTGTACAGAA 1719
 Db 491 LeuGlnValIlePheSerValThrValLeuCysGlnLysLeuMetGlnValProGln 510
 QY 1720 GAA 1722
 Db 511 Glu 511
 RESULT 5
 AAY58426
 ID AAY58426 standard; Protein; 512 AA.
 XX

AC AAY58426;
 XX 27-MAR-2000 (first entry)
 DT Rat neutral amino acid transporter LAT1.
 XX Rat; neutral amino acid transporter; LAT1; drug permeation.
 KW Rattus sp.
 OS JP11299489-A.
 PN 02-NOV-1999.
 XX 22-APR-1998; 98JP-0126648.
 XX 22-APR-1998; 98JP-0126648.
 PR (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 XX WPI: 2000-075280/07.
 XX N-PSDB; AAY55789.
 DR A neutral amino acid transporter and its gene -
 XX Claim 1; Page 9; 17pp; Japanese.
 PS This sequence represents a rat neutral amino acid transporter,
 CC LAT1. It interacts with a 528 amino acid protein (AAY58427)
 CC to effect neutral amino acid transport. LAT1 is predicted to have up to
 CC twelve transmembrane domains from hydropathy analysis. The protein, and
 CC nucleic acids encoding it are useful for the in vitro analysis of
 CC drug permeation through cell membranes and blood tissue barriers.
 XX SQ Sequence 512 AA;
 Alignment Scores:
 Pred. No.: 6.22e-127 Length: 512
 Score: 1204.00 Matches: 229
 Percent Similarity: 67.57% Conservative: 96
 Best Local Similarity: 47.61% Mismatches: 150
 Query Match: 29.22% Indels: 6
 DB: 2 Gaps: 2
 US-09-667-170A-440 (1-2239) x AAY58426 (1-512)
 QY 298 CACGCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGACAGAGAA 357
 Db 31 GluAlaArgArgGlyAspGlyAlaAspProGluGlyGlyValThrLeuGlnArgAsn 50
 QY 358 GTCACCTTTACTGAGGGAGTCTCCATTATTCGACCATTCATTGGAGGAGGAATCTTC 417
 Db 51 IleThrLeuIleAsnGlyValAlaIleIleValGlyThrIleIleGlySerGlyIlePhe 70
 QY 418 ATCTCTCTAAAGGGCTGCTCCAGAACACGCGGCGAGCTGCTCTGACCATCTGG 477
 Db 71 ValThrProThrGlyValLeuLysGluAlaGlySerProGlyLeuSerLeuValTrp 90
 QY 478 ACGGTGTGTGGGCTCCTCTACTATTTCGAGCTTGTCTTATGCTGATTCGGAGCAACT 537
 Db 91 AlaValCysGlyValPheSerIleValGlyAlaLeuCysTyrrAlaGluLeuGlyThrThr 110
 QY 538 ATAAAGAAATCGAGGTCATTACACATATATTTTGGAAAGTCTTTGGTCCATTACCAGCT 597
 Db 111 IleSerLysSerGlyAspTyrrAlaTyrrMetLeuGluValTyrrGlySerLeuProAla 130
 QY 598 TTTGTACGAGTCTGGGTGGAACCTCTCATATACGCCCTGCAGCTACTGCTGTGATATCC 657
 Db 131 PheLeuLysLeuTyrrIleGluLeuIleIleArgProSerSerGlnTyrrIleValAla 150
 QY 658 CTGGCATTTGGAGCGCTACATCTCGAAGCAATTTTTTATTAATGTGAATCCCTGAACCT 717
 Db 151 LeuValPheAlaThrTyrrLeuLeuLysProValPheProThrCysProValProGlu 170

QY 718 GCGATCAAGCTCATTACAGCTGCGGCATACCTAGTGTGCTTAAATAGCATGAGT 777
 DB 171 AlalaLysLeuValAlaCysLeuValLeuLeuThrAlaValAsnCysTyrSer 190
 QY 778 GTCAGCTGGAGCGCCGAGTCCAGATTTTCTTAACCTTTTGCAGCTCACAGCAATCTG 837
 DB 191 ValLysAlaAlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAla 210
 QY 838 ATAATATATAGTCCGAGTATGAGCTAATTTAA-----GGTCAACCCAG 885
 DB 211 LeuLeuLeuLeuGlyPheLeuGlnMetGlyLysAspLeuGlyGlnGlyAspAlaSer 230
 QY 886 AACTTT-----AAAGAGCGCTTTTTCAGGAAGAGATCAAGTATTACCGGTGCCACTG 939
 DB 231 AsnLeuHisGlnLysLeuSerPheGluGlyThrAsnLeuAspValGlyAsnLeuValLeu 250
 QY 940 GCTTTTATTATGGAATGATGATGATGCTGCTGGTGTATACCTCAACTTTGTACTGAA 999
 DB 251 AlaLeuTyrSerGlyLeuPheAlaTyrGlyTyrAsnTyrLeuAsnPheValThrGlu 270
 QY 1000 GAAGTAGAAACCCCTGAAAAACCATTCCTTCATATGATATATCCATGGCCATGTC 1059
 DB 271 GluMetIleAsnProTyrArgAsnLeuProLeuAlaIleLeuSerLeuProIleVal 290
 QY 1060 ACCATTGGCTATGCTGACAAATGTGGCTTACTTTACGACCAATTAATGCTGAGGAGCTG 1119
 DB 291 ThrLeuValTyrValLeuThrAsnLeuAlaTyrPheThrThrLeuSerThrAsnGlnMet 310
 QY 1120 CTGCTTCAATGAGTGGGAGTACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTA 1179
 DB 311 LeuThrSerGluAlaValAlaValAspPheGlyAsnTyrHisLeuGlyValMetSerTyr 330
 QY 1180 GCAGTTCGAGCTTTGTTGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1239
 DB 331 IleIleProValPheValGlyLeuSerCysPheGlySerValAsnGlySerLeuPheThr 350
 QY 1240 GTCCTCAGGTTATTTATGTTGGCTTCGAGAGGCTACCTTCCAGAAATTTCTCCATG 1299
 DB 351 SerSerArgLeuPhePheValGlySerArgGluGlyHisLeuProSerIleLeuSerMet 370
 QY 1300 ATTATGTCGCGAGCAGCTCTCTACAGCTGTTATTTGTTGTCACCTTTTGCAATG 1359
 DB 371 IleHisProGlnLeuLeuThrProValProSerLeuValPheThrCysValMetThrLeu 390
 QY 1360 ATATGCTCTCTCTGAGAGCTCGAGAGCTTTTGAATTTCTCTCAGTTTGCAGGTGG 1419
 DB 391 MetTyrAlaPheSerArgAspPhePheSerIleIleAsnPhePheSerPheAsnTyr 410
 QY 1420 CTTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1479
 DB 411 LeuCysValAlaLeuAlaIleIleGlyMetMetTrpLeuArgPheLysLysProGluLeu 430
 QY 1480 CATGCTCTCTTCAAGTGGCCATGTTTCATCCAGCTTTTCTTCTTCTCATGCTCTCTC 1539
 DB 431 GluArgProIleLysValAsnLeuAlaLeuProValPhePheLeuLeuAlaCysLeuPhe 450
 QY 1540 ATGCTTCCCTTCCCTCTATTCCGACCCATTTAGTACAGGATGCTGCTGCTGCTGCTGCT 1599
 DB 451 LeuIleAlaValSerPheTyrPheThrProLeuGluCysGlyIleGlyPheAlaIle 470
 QY 1600 CTGACTGGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1659
 DB 471 LeuSerGlyLeuProValTyrPhePheGlyValTyrTrpLysAsnLysProLysTrpIle 490
 QY 1660 AGAATATGTCAGAGAAATACCAACACATTAATTAATTAATTAATTAATTAATTAATTAAT 1719
 DB 491 LeuGlnValIlePheSerValThrValLeuCysGlnLysLeuMetGlnValValProGln 510
 QY 1720 GAA 1722
 DB 511 Glu 511

RESULT 6
 ABB12114
 ID ABB12114 standard; peptide; 507 AA.
 XX
 AC ABB12114;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human 4F2 light chain homologue, SEQ ID NO:2484.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cyostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnerary; antiulcer.
 XX
 OS Homo sapiens.
 XX
 XX WO200157188-A2.
 PN
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-457740/49.
 DR N-PSDB; ABA09358.
 DR
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 PS
 XX
 XX Claim 20; Page 308-309; 1963pp; English.
 CC
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and

CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.

SQ Sequence 507 AA:

Alignment Scores: 1.14e-119 Length: 507
 Pred. No.: 1140.00 Matches: 218
 Score: 66.17% Conservative: 95
 Percent Similarity: 46.09% Mismatches: 156
 Best Local Similarity: 27.67% Indels: 4
 Query Match: 22 Gaps: 3
 DB:

US-09-667-170A-440 (1-2239) x ABB12114 (1-507)

QY 247 CTTCTCCCTGGCCCAACAGGACCCAGCTGGGCGAGCGCCTTTTCAGGAAGAGACGCGCTTT 306
 DB 14 ProValAlaGluGluGlyGluGluAlaArgGluGlyIleMetAlaAlaLysArgAlaAsp 33
 QY 307 TCAGGAAGAGACGCGCTTTTCAGGAAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTA 366
 DB 34 GlyAlaAlaProAlaGlyGluGly---GluGlyValThrLeuGlnGlyAsnIleThrLeu 52
 QY 367 CTGAGGGAGCTCCATTATTCATTCATGCGACCATCATGAGGAGGAAATCTCATCTCTCT 426
 DB 53 LeuLysGlyValAlaValIleValAlaIleValAlaIleMetGlySerGlyIlePheValThrPro 72
 QY 427 AAGGGCTGTCTCCAGAACACGCGCGCTGGCATCTCTCTGACCATCTGGAGCGGTGTCT 486
 DB 73 ThrGlyValLeuLysGluAlaGlySerProGlyLeuAlaLeuValValTrpAlaAlaCys 92
 QY 487 GGGTCTCTGCTACATTTGAGGCTGTCTATGCTGAATGGGAAACAACTATAAGAA 546
 DB 93 GlyValPheSerIleValGlyAlaLeuGlyCysValAlaGluLeuGlyThrIleSerLys 112
 QY 547 TCTGGAGTCAATACATATATTTTGGAGTCTTTGGTCCATTACAGCTTTTGTACGA 606
 DB 113 SerGlyGlyAspTyrAlaTyrMetLeuAspValTyrGlySerLeuProAlaPheLeuLys 132
 QY 607 GTCTGGTGGACCTCCATTAATACGCGCTGACCTACTCTGTGATATCCCTGGCATTT 666
 DB 133 LeuTrpIleGluLeuLeuIleIleArgProSerSerGlnTyrIleValAlaLeuValPhe 152
 QY 667 GGACGCTACATCTCGAACCATTTTATTCAATGAAATCCCTGAATCTGGGATCAAG 726
 DB 153 AlaThrTyrLeuLeuLysProLeuPheProThrCysProValProGluGluAlaAlaLys 172
 QY 727 CTCATTACAGCTGTGGCATACTAGTAGTATGCTTAATAGCATGAGTGTACAGCTGG 786
 DB 173 LeuValAlaCysLeuGlyValLeuLeuLeuThrAlaValAsnCysTyrSerValLysAla 192
 QY 787 AGCGCCGGATCCAGATTTCTTAACCTTTTGGAGCTCACACAAATCTGATAAATATA 846
 DB 193 AlaThrArgValGlnAspAlaPheAlaAlaLysLeuLeuAlaLeuAlaLeuIle 212
 QY 847 GTCCCTGGAGTATGACAGCTAAATTAAGGTCACAAACGAGCAACTTTAAAGACGCTTTTCA 906
 DB 213 LeuLeuGlyPheValGlnIleGlyLysGlyAspValSerAsnLeuAspProAsnPheSer 232
 QY 907 -----GGAAGAGATTCAGATATTACGGGTGTGCGCTGCTTTTATATGGAATGAT 960
 DB 233 PheGluGlyThrLysLeuAspValGlyAsnIleValLeuAlaLeuTyrSerGlyLeuPhe 252

QY 961 GCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCCCTGAAAA 1020
 DB 253 AlaTyrGlyGlyTrpAsnTyrLeuAsnPheValThrGluGluMetIleAsnProTyrArg 272
 QY 1021 ACCATTCCCTTGAATATGATATATCATCGCCATTCACCATTCGCTATGCTGCTGACA 1080
 DB 273 AsnLeuProLeuAlaIleIleSerLeuProIleValThrLeuValTyrValLeuThr 292
 QY 1081 AATCGCTCTACTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAATAGCTGGCA 1140
 DB 293 AsnLeuAlaTyrPheThrThrLeuSerThrGluGlnMetLeuSerSerGluAlaValAla 312
 QY 1141 GTGACCTTTTCTGAGCGCTACTGGGAAATTTCTCATAGCAGTTCGATCTTTGTTGCC 1200
 DB 313 ValAspPheGlyAsnTyrHisLeuGlyValMetSerTrpIleIleProValPheValGly 332
 QY 1201 CTCCTCTGCTTTGCTCATCAAGCGGTGGTGTCTGCTCTCCAGGTTATCTATGTT 1260
 DB 333 LeuSerCysPheGlySerValAsnGlySerLeuPheThrSerSerArgLeuPhePheVal 352
 QY 1261 CGCTCTCGAGAGGCTCACTTCCAGAAATCTCTCCATGATTCATGTCGCCAAGCACACT 1320
 DB 353 GlySerArgGluGlyHisLeuProSerIleLeuSerMetIleHisProGlnLeuLeuThr 372
 QY 1321 CCTCTACAGCTGTATTGTTTGCACCTTTGCACATGATATGCTCTCTCTGGAGAC 1380
 DB 373 ProValProSerLeuValPheThrCysValMetThrLeuPheTyrAlaPheSerLysAsp 392
 QY 1381 CTCGACAGCTTTTGAATTTCTCAGTTTTCAGGTTGGCTTTTATTGGCTGGCAGTT 1440
 DB 393 IlePheSerValIleAsnPhePheSerPhePheAsnTrpLeuCysValAlaLeuAlaIle 412
 QY 1441 GCTGGCTGATTTATCTTCGATCAAAATCCCTCCATGATTCATGTCGCCAAGTGCCA 1500
 DB 413 IleGlyMetIleTrpLeuArgHisArgLysProGluLeuGluArgProIleLysValAsn 432
 QY 1501 CTGTTTCATCCAGCTTTGTTTCTTCATGCGCTCTTCATGTTGCGCTTCCCTCTCTAT 1560
 DB 433 LeuAlaLeuProValPhePheIleLeuAlaCysLeuPheLeuIleAlaValSerPheTrp 452
 QY 1561 ---TCGGACCCCATTTAGTACAGGATGGCTTCGTCATCTCACTCTGAGTGGAGTCCCTGGC 1617
 DB 453 LysThrThrProTyrSerValAlaSerAspPheThrIleIleLeuSerGlyLeuProVal 472
 QY 1618 TATTATCTCTTTATATATATGAGGACAAAGAACCCAGGTGG 1656
 DB 473 TyrPheGlyValTrpTrpLysAsnLysProLysTrp 485

RESULT 7
 AAY41159
 ID AAY41159 standard; Protein; 511 AA.
 XX
 AC AAY41159;
 XX
 DT 24-JAN-2000 (first entry)
 XX
 DE Human amino acid permease homolog (AAPH).
 XX
 KW Amino acid permease homolog; AAPH; cancer; inflammatory disorder; AIDS;
 KW autoimmune disorder; anemia; asthma; atherosclerosis; diabetes mellitus;
 KW multiple sclerosis; osteoarthritis; osteoporosis; rheumatoid arthritis;
 KW ulcerative colitis; infection; cell proliferation disorder; human.
 OS Homo sapiens.
 XX
 PN US5981242-A.
 XX
 PD 09-NOV-1999.
 XX
 PF 03-FEB-1999; 98US-0243920.
 XX
 PR 06-MAY-1999; 98US-0073362.
 XX

(INCY-) INCYTE PHARM INC.

Yue H, Corley NC, Hillman JL;

WPI; 1999-633325/54.

N-PSDB; AAZ23240.

New human amino acid permease homolog, useful in the diagnosis, treatment and prevention of cancer, inflammatory/autoimmune disorders and cell proliferation disorders

Claim 1; Fig 1A-F; 32pp; English.

This represents a human amino acid permease homolog (AAPH). The AAPH nucleic acid, protein, AAPH (antagonists are useful in the diagnosis, treatment and prevention of cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid or uterus; inflammatory/autoimmune disorders such as AIDS, Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, diabetes mellitus, Crohn's disease, atopic dermatitis, dermatomyositis, emphysema, multiple sclerosis, osteoarthritis, osteoporosis, rheumatoid arthritis, ulcerative colitis, viral, bacterial, fungal, parasitic, protozoal infections; and cell proliferation disorders such as actinic, keratosis, cirrhosis, arteriosclerosis, myelofibrosis, and primary thrombocythemia.

SQ Sequence 511 AA;

Alignment Scores:

Pred. No.: 1,05e-118 Length: 511
Score: 1131.50 Matches: 214
Percent Similarity: 68.00% Conservative: 109
Best Local Similarity: 45.05% Mismatches: 149
Query Match: 27.46% Indels: 3
DB: 20 Gaps: 2

US-09-667-170A-440 (1-2239) x AAY41159 (1-511)

QY 310 GGAAGAGACGCTTTTTCAGGAGAGAGAAAGTCAGCTGAAGAGAAAGTCACTTTACTG 369
Db 21 GlyAspGlyAlaSerProGlyProGluInValLysLeuLysLysLysLysLysLeuLeu 40
QY 370 AGGGAGCTCCATTATCATTCATGGCACCATCATTTGGAGCAGGAATCTCTCTCTAAG 429
Db 41 AsnGlyValCysLeuLeuValGlyAsnMetIleGlySerGlyIlePheValSerProLys 60
QY 430 GGCGTGTCTCCAGAACGCGCGGCGGATCTCTGACCATCTGACCATCTGGAGGTGTGGG 489
Db 61 GlyValLeuLeuTyrSerAlaSerPheGlyLeuSerLeuValIleTyrPheAlaValGlyGly 80
QY 490 GTCTGTCTACATTTTGGAGCTTTTCTTCTGTAATGGGAACTATAAGAAATCT 549
Db 81 LeuPheSerValPheGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleLysLysSer 100
QY 550 GGAGGTCATTACACATATATTTTGGAGCTTTTGGTCCATTACAGCTTTTGTACGAGTC 609
Db 101 GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheIleArgLeu 120
QY 610 TGGGTGGAATCTCTCATATACGCTGCGAGCTACTCTGTGATATCCCTGGCATTTTGA 669
Db 121 TrpThrSerLeuLeuIleIleGluProThrSerGlnAlaIleIleAlaIleThrPheAla 140
QY 670 CGCTACATCTGGAACCATTTTATTTCATATGTGAATCCCTGAACCTTGCAGTCAAGCTC 729
Db 141 AsnTyrMetValGlnProLeuPheProSerCysPheAlaProTyrAlaAlaSerArgLeu 160
QY 730 ATTACAGCTGTGGCATACTGTAGTGTGGTCTTAATAGCATGAGTCTCAGCTGGAGC 789

Db 161 LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyrValLysTrpGly 180
QY 790 GCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATCTGTATAATATATAGTC 849
Db 181 ThrLeuValGlnAspIlePheThrTyrAlaLysValLeuAlaLeuIleAlaValIleVal 200
QY 850 CTGGAGTTATCCAGCTAATTAAGGTCAACGCAGAACTTTAAAGACCCCTTTTCAGGA 909
Db 201 AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGluGly 220
QY 910 AGAGATTCAAGTATTACCGCGTTGCCACTGGCTTTTATTATGGAATGATCATATGCT 969
Db 221 SerSerPheAlaValGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer 240
QY 970 GGCTGGTTTACTCACTCACTTGTACTGAAGAGTAGAAACCCCTCAAAAACCATTCCTC 1029
Db 241 GlyTyrAspThrLeuAsnTyrValThrGluGluIleLysAsnProGluArgAsnLeuPro 260
QY 1030 CTGCAATATGATATCCATGGCCATTTGCACCATTTGGCTATGCTGCAAAATGTGGCC 1089
Db 261 LeuSerIleGlyLeuSerMetProIleValThrIleIleIleTyrIleLeuThrAsnValAla 280
QY 1090 TACTTTACGACCATTAATGCTGAGGAGCTGCTTTTCAAAATGAGTGGCAGTGCACCTTT 1149
Db 281 TyrTyrThrValLeuAspMetArgaspIleLeuAlaSerAspAlaValAlaValThrPhe 300
QY 1150 TCTGAGCGCTACTGGGAAATTTCTCATPAGCAGTTCGATCTTTGTTGCCCTCTCTGTC 1209
Db 301 AlaAspGlnIlePheGlyIlePheAsnTrpIleIleProLeuSerValAlaLeuSerCys 320
QY 1210 TTTGGCTCCATGAACGGTGTGTTGCTGCTCCAGGTTATTTATGTTGCGTCTCGA 1269
Db 321 PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuPhePheValGlySerArg 340
QY 1270 GAGGTCACCTTCCAGAAATCTCTCATGATTCATGTCGCGAGCACACCTCTCTACCA 1329
Db 341 GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro 360
QY 1330 GCTGTTATTTGTTGACCCCTTTGACATGATGATGCTCTCTCTGGAGACCTCGACAGT 1389
Db 361 SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln 380
QY 1390 CTTTGAATTTCTCAGTTTCCAGGTGGCTTTTATTTATGCGTGGCAGTCTCTGGCTG 1449
Db 381 LeuIleAsnTyrTyrSerPheSerTyrTrpPhePheValGlyLeuSerIleValGlyGln 400
QY 1450 ATTTATCTTCGATACAAATGCCAGATATGTCATCTCTCTTCAAGTGGCCTGTTCTATC 1509
Db 401 LeuTyrLeuArgTyrLysGluProAspArgProArgProLeuLysLeuSerValPhePhe 420
QY 1510 CCAGCTTTGTTTCTTCTTCCATGCTCTTCATGGTGGCTTTCCTCTATTCGAGACCA 1569
Db 421 ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr 440
QY 1570 TTTTACGAGGATTTGGCTTCTGTCATCACTCTGAGTGGCTCCCTGCTATTCCTCTT 1629
Db 441 IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuPhePheTyrPheLeuIle 460
QY 1630 ATT-----ATATGGGACAAAGAACCCAGGTGGTTTGAATAATTCAGAGAAAAATAACC 1683
Db 461 IleArgValProGluHisLysArgProLeuTyrLeuArgArgIleValGlySerAlaThr 480
QY 1684 AGAACATTACAAATAATA---CTGGAAGTTGTACACAGAGAGAT 1725
Db 481 ArgTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetAsp 495
RESULT 8
AAY27076
ID AAY27076 standard; Protein; 511 AA.
XX
AC AAY27076;
XX

18-OCT-1999 (first entry)
 Human amino acid permease homologue (AAPH).
 Amino acid permease homologue; AAPH; cancer; inflammatory disorder;
 autoimmune disorder; cell proliferation disorder; adenocarcinoma;
 AIDS; Crohn's disease; rheumatoid arthritis; arteriosclerosis; human;
 psoriasis; hepatitis; gene therapy.
 Homo sapiens.
 US5942399-A.
 24-AUG-1999.
 06-MAY-1998; 98US-0073362.
 06-MAY-1998; 98US-0073362.
 (INCY-) INCYTE PHARM INC.
 Corley NC, Hillman JL, Yue H;
 WPI; 1999-493508/41.
 N-PSDB; AAX86999.
 New human amino acid permease homolog (AAPH) polynucleotide and its
 complement useful for prevention and treatment of cancer, autoimmune
 and cell proliferation disorders
 Claim 1; Fig 1A-F; 33pp; English.
 This sequence represents a human amino acid permease homologue (AAPH).
 The protein can be expressed by standard recombinant methodology. AAPH
 polynucleotides and polypeptides are useful for diagnosis, treatment and
 prevention of cancer e.g. adenocarcinoma, lung cancer, ovarian cancer;
 inflammatory/autoimmune disorders e.g. AIDS, Crohn's disease, rheumatoid
 arthritis; and cell proliferation disorders e.g. arteriosclerosis,
 psoriasis, hepatitis. AAPH polynucleotides are useful in gene therapy,
 in PCR-based methods for detecting upstream sequences, and also for
 generating hybridisation probes for mapping the naturally occurring
 genomic sequence.
 SQ Sequence 511 AA;
 Alignment Scores:
 Pred No.: 1 05e-118 Length: 511
 Score: 1131.50 Matches: 214
 Percent Similarity: 68.00% Conservative: 109
 Best Local Similarity: 45.05% Mismatches: 149
 Query Match: 27.46% Indels: 3
 DB: 2 Gaps: 2
 US-09-667-170A-440 (1-2239) x AAY27076 (1-511)
 QY 310 GGAAGACAGCGCTTTTCAGGAGAGAGAGAGTGCAGCTGAAGAGAGATCCTTACTG 369
 DB 21 GlyAspGlyAlaSerProGlyProGluGlnValLysLeuLysLysGluLysSerLeuLeu 40
 QY 370 AGGGAGTCTCCATTATCATTTGGCAGCACCATTATGGAGCAGCAATCTCTCTCTAAG 429
 DB 41 AsnGlyValCysLeuValGlyValGlyAsnMetileGlySerGlyIlePheValSerProLys 60
 QY 430 GCGGTCTCCAGACAGCGGCGGCGATGTCTCTGACCATCTGTGGAGCGGTGTGGG 489
 DB 61 GlyValLeuLeuIleTyrSerAlaSerPheGlyLeuSerLeuValIleTyrAlaValGlyGly 80
 QY 490 GTCCTGTCTACTATTGGAGCTTTGTCTTATGCTGAATTTGGAGACAACTATAAGAAATCT 549
 DB 81 LeuPheSerValPheGlyAlaLeuCysTyrAlaGluLeuGlyThrThrIleLysLysSer 100
 QY 550 GGAGGTCTATACATATATTTTGGAACTCTTTGGTCCATTACCAGCTTTTGTACAGTCT 609
 DB 101 GlyAlaSerTyrAlaTyrIleLeuGluAlaPheGlyGlyPheLeuAlaPheLeuArgLeu 120
 QY 610 TGGGTGGAACCTCTCATAATACGCGCTGAGCTACTCTGTGATATCCCTGGCATTTTGA 669
 DB 121 TrpThrSerLeuLeuIleIleGluProThrSerGlnAlaIleIleAlaIleThrPheAla 140
 QY 670 CGCTACATCTCTGGAACCACTTTTATTTCAATGTGAATCCCTGAACCTTGGAGTCACTC 729
 DB 141 AsnTyrMetValGlnProLeuPheProSerCysPheAlaProTyrAlaAlaSerArgLeu 160
 QY 730 ATTACAGCTGTGGCATAACTGTAGTGTCTTAATAGCATGAGTGTGAGTGTGGAGC 789
 DB 161 LeuAlaAlaAlaCysIleCysLeuLeuThrPheIleAsnCysAlaTyrValIleTyrPgly 180
 QY 790 GCGCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATCTCTGATAATATATAGTC 849
 DB 181 ThrLeuValGlnAspIlePheThrTyrAlaLysValLeuAlaLeuIleAlaValIleVal 200
 QY 850 CTGGAGTTATGCAGCTAAATTAAGGTCACACGAGCAACTTTTAAAGACGCTTTTTCAGGA 909
 DB 201 AlaGlyIleValArgLeuGlyGlnGlyAlaSerThrHisPheGluAsnSerPheGluGly 220
 QY 910 AGAGATTCAAGTATTACGCGGCTTCCACTGCTTTTATTATGGAATGTATGATATGCT 969
 DB 221 SerSerPheAlaValGlyAspIleAlaLeuAlaLeuTyrSerAlaLeuPheSerTyrSer 240
 QY 970 GGCTGTTTACCTCACTTTTGTACTGAGAGAGTAGAAGAACCTGAAACCAATTCCTCC 1029
 DB 241 GlyTrpAspThrLeuAsnTyrValThrGluIleLysAsnProGluArgAsnLeuPro 260
 QY 1030 CTTCGAATATGATATCATGCGCATTTGTCACCATTTGCTGATGTGTGACAAATGTGGCC 1089
 DB 261 LeuSerIleGlyIleSerMetProIleValThrIleIleTyrIleLeuThrAsnValAla 280
 QY 1090 TACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAATGAGTGGAGTGCACCTTT 1149
 DB 281 TyrTyrThrValLeuAspMetArgAspIleLeuAlaSerAspAlaValAlaValThrPhe 300
 QY 1150 TCTGAGCGCTACTGGAATTTCTCATTAGCAGTCCGATCTTTGTTGGCCCTCTCTGTC 1209
 DB 301 AlaAspGlnIlePheGlyIlePheAsnTrpIleIleProLeuSerValAlaLeuSerCys 320
 QY 1210 TTTGGCTCCATGACGCGGTGTGTGCTGCTCTCCAGGTTATTTCTATGTTGCTGCTCGA 1269
 DB 321 PheGlyGlyLeuAsnAlaSerIleValAlaAlaSerArgLeuPheValGlySerArg 340
 QY 1270 GAGGTCCACCTCCAGAAATCTCTCCATGATTCATGTCGCGACACACCTCTCTACCA 1329
 DB 341 GluGlyHisLeuProAspAlaIleCysMetIleHisValGluArgPheThrProValPro 360
 QY 1330 GCTGTTATGTTTTCACCCCTTTGACATGATATGCTCTCTCTGAGAGCTCGACAGT 1389
 DB 361 SerLeuLeuPheAsnGlyIleMetAlaLeuIleTyrLeuCysValGluAspIlePheGln 380
 QY 1390 CTTTGAATTTCTCAGTTTCCAGGTGGCTTTTATTTGGCTGGCAGTGTGGGCTG 1449
 DB 381 LeuIleAsnTyrTyrSerPheSerTyrTrpPhePheValGlyLeuSerIleValGlyGln 400
 QY 1450 ATTTATCTTCGATACAAATGCGCAGATATGATGCTCTCTTCAAGGTGCCACTGTTCATC 1509
 DB 401 LeuTyrLeuArgTrpLysGluProAspArgProArgProLeuLysLeuSerValPhePhe 420
 QY 1510 CGAGTGTGTTTTCCTTCACATGCTCTTCATGTTGGTGTGGCTTCCCTCTATTCGAGCCA 1569
 DB 421 ProIleValPheCysLeuCysThrIlePheLeuValAlaValProLeuTyrSerAspThr 440
 QY 1570 TTTAGTACAGGATTTGGCTTCTCATCACTCTGAGTGGAGTCCCTGCTATTTATCTCTTT 1629
 DB 441 IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheTyrPheLeuIle 460
 QY 1630 ATTT-----ATATGGACAAAGAACCCAGGTGGTGTAGATAATATGTCAGAGAAATAACC 1683
 DB 461 IleArgValProGluHisLysArgProLeuTyrLeuArgGluIleValGlySerAlaThr 480

QY 1570 TTATGACAGGAGGATTGGCTTGGTCACTCTGAGTCCCTGGCTGATATATCTCTTT 1629
 Db 441 IleAsnSerLeuIleGlyIleAlaIleAlaLeuSerGlyLeuProPheThrPheLeuIle 460
 QY 1630 ATT-----ATATGGACAGAAACCCAGTGGTTTAGAATAATGTCAGAGAAATAACC 1683
 Db 461 IleArgValProGluHisLysArgProLeuTyrLeuArgIleValGlySerAlaThr 480
 QY 1684 AGAACATTACAAATAATA---CTGGAAGTTGTACCAAGAAAGAT 1725
 Db 481 ArgTyrLeuGlnValLeuCysMetSerValAlaAlaGluMetAsp 495

RESULT 10
 ABB58292
 ID ABB58292 standard; Protein; 505 AA.
 XX
 AC ABB58292;
 DE
 DF 26-MAR-2002 (first entry)
 XX Drosophila melanogaster polypeptide SEQ ID NO 1668.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL02395.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 1668; 2lpp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 505 AA;

Alignment Scores:
 Pred. No.: 1.16e-116 Length: 505
 Score: 1113.50 Matches: 208
 Percent Similarity: 64.50% Conservative: 110
 Best Local Similarity: 42.19% Mismatches: 170
 Query Match: 27.03% Indels: 5
 DB: 22 Gaps: 2

US-09-667-170A-440 (1-2239) x ABB58292 (1-505)

QY 235 AACGGGAGGTGCTTCCCTGGGCAACAAGAGCCACCTGGGAGGAGCGCTTTTCAGGA 294
 Db 11 SerSerArgLysSerSerIleValAsnGlyAsnGlyAspAlaSerAlaLysLeuThrAsn 30
 QY 295 AGAGACGCCCTTTTCAGGAGAGAGCGCTTTTCAGGAGAGAGAGAAAGTCAGCTCAGAGG 354
 Db 31 GlyAspGlyAspGlyGlyAsp-----GlyGlyGlyGluValThrLeuLysAla 47
 QY 355 AAATCCTACTTACTGAGGGAGTCTCCATATATATTTGGACCATCATTTGGAGCAGGAATC 414
 Db 48 LysMetSerLeuLeuAsnGlyCysThrValIleValGlySerIleIleGlySerGlyIle 67
 QY 415 TTCATCTCTCCTAAAGGCGTCTCCAGAACACGGCGAGCGTGGGCATGCTCTGCACATC 474
 Db 68 PheValSerProThrGlyValLeuMetTyrThrGlySerValAsnLeuAlaIleVal 87
 QY 475 TGGAGGTGTGTGGGCTCTGCTCACTATTGGAGCTTTGTCTTATGCTGAATGGGAACA 534
 Db 88 TrpValIleSerGlyLeuPheSerMetValGlyAlaTyrCysTyrAlaGluLeuGlyThr 107
 QY 535 ACTATAAGAAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTTGGTCCATTACCA 594
 Db 108 MetIleThrLysSerGlyAlaAspTyrAlaTyrIleMetGluThrPheGlyProPheMet 127
 QY 595 GCTTTGTACGAGTCTGGTGGAACTCTCTATATACGCCCTGCAGCTACTGCTGTGATA 654
 Db 128 AlaPheIleArgLeuTrpIleGluCysMetIleValArgProCysSerGlnAlaIleVal 147
 QY 655 TCCCTGGCATTTGGACGCTACATCTGGAACCATTTTTTATTCATGTAATCCCTGAA 714
 Db 148 AlaLeuThrPheSerThrTyrValLeuLysProPhePheProGluCysThrProGlu 167
 QY 715 CTTGCGATCAAGCTCATTCAGCTGTGGCATTAACCTAGTGTAGTGGTCTCTAAATAGCATG 774
 Db 168 AspSerAlaArgLeuAlaValCysCysIleLeuValLeuThrLeuIleAsnCysTrp 187
 QY 775 AGTGTCAGTGGAGCGCGGATCCAGATTTCCTTAACCTTTTGCAGCTCACAGCAATT 834
 Db 188 AspValLysTrpAlaThrAlaValGlnAspIlePheThrTyrAlaLysLeuAlaLeu 207
 QY 835 CTGATAATTATAGTCCCTGGAGTTATGAGCTAATTAAGTCAACACGAGCACTTTAAA 894
 Db 208 PheIleIleAlaThrGlyValTyrGlnLeuTyrLeuGlyAsnThrGlnTyrPhe--- 226
 QY 895 GAGCGCTTTTCAGGAGAGATTCAAGTATTACGCGGTGGCCACTGCTTTTATTTATGGA 954
 Db 227 ---ThrPheGluAsnThrAspThrLysValThrSerIleAlaLeuSerPheTyrSerGly 245
 QY 955 ATGTATCATATGCTGGCTGGTGTTCCTCACTCACTTGTGTGTAAGAGAGTACAAACCT 1014
 Db 246 LeuPheAlaTyrAsnGlyTrpAsnTyrLeuAsnPheIleIleGluGluLeuLysPro 265
 QY 1015 GAAACCACTTCCCTGCAATATATATCCATGTCATGTCACCATTTGCTATGATG 1074
 Db 266 ValLysAsnLeuProArgAlaIleAlaIleSerCysThrLeuValThrIleValTyrVal 285
 QY 1075 TGCACAAATGCGCTTACTTTAGCACCATTAATGCTGAGGAGCTGCTGCTTCAATGCA 1134
 Db 286 MetAlaAsnValSerPheTyrThrIleLeuSerProAspGluValMetGlySerAla 305
 QY 1135 GTGGAGTACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCACTTCCGATCTTT 1194
 Db 306 ValAlaValThrTyrAlaGluArgAlaPheGlyMeLeuAlaThrThrIleProValPhe 325
 QY 1195 GTTGCCCTCTCTGCTTGGCTCCATGAACGGTGGTGTGTGCTGCTCTCCAGGTATTTC 1254
 Db 326 ValAlaLeuSerThrPheGlyAlaValAsnGlyIleLeuLeuThrSerSerArgLeuPhe 345
 QY 1255 TATGTTGCGTCTCGAGGCGTACCTTCCAGAAATCTCTCCAGATGATTATGCTCCCAAG 1314
 Db 346 TyrAlaGlyAlaAsnAsnGlyGlnMetProGluIleLeuThrMetIleGlnIleArg 365
 QY 1315 CACACTCTCTACCAGCTGTTATTGTTTTCACCCCTTTTGACATGATATGCTCTCTCTCT 1374

Db 366 PheThrProThrProAlaValLeuAlaMetAlaLeuSerMetLeuTyrLeuThrVal 385
 QY 1375 GGAGACCTCGACAGTCTTTGAATTCCTCAGTTTCCAGGTGGCTTTTATTGGCTG 1434
 Db 386 SerAspIlePheAlaLeuIleAsnTyrValGlyPheAlaThrTrpLeuSerIleGlyVal 405
 QY 1435 CGAGTTGCTGGCTGATTATCTTCGATACAAATGCCAGATATGCATGCTCTTCAAG 1494
 Db 406 AlaValLeuCysLeuProThrPheAlaThrTrpAlaGlnProAsnLeuProArgProIleArg 425
 QY 1495 GTGCACTGTTTCATCCAGCTTTTCTTCCTTCATGCTCTTCATGCTGGTGGCTTCC 1554
 Db 436 ValProMetValPheProIleValTyrLeuIleAlaThrIlePheValThrValValPro 445
 QY 1555 CTCATTCGGACCCATTTAGTACAGGATGGCTTCCTCATCTCATCTGCTGAGTCCCT 1614
 Db 446 MetTyrAlaSerProValGluThrGlyTyrGlyIleLeuMetIleLeuSerSerIlePro 465
 QY 1615 CGTATTATCTCTTTATATATGGCAAGAAACCCAGGTGGTTCAGAAATAGTCAGAG 1674
 Db 466 ValTyrLeuValPheIleAlaTrpLysAsnLysProIleThrPheGlnLysThrMetGly 485
 QY 1675 AAATACCCAGACATTTACAAATATATCTGGAAGTCTA 1713
 Db 486 GlyLeuThrGlnValLeuGlnLysLeuMetMetValVal 498
 RESULT 11
 AAB72393
 ID AAB72393 standard; Protein; 533 AA.
 AC AAB72393;
 XX
 DT 24-MAY-2001 (first entry)
 XX
 DE Rat L-type amino acid transporter 2 (LAT2) amino acid sequence.
 XX
 KW Neutral amino acid transporter; amino acid transport activating factor;
 4F2hc; 4F2 heavy chain; drug permeation; poison; rat; LAT2;
 XX L-type amino acid transporter 2.
 XX
 OS Rattus sp.
 XX
 PN JP2000342270-A.
 XX
 PD 12-DEC-2000.
 XX
 PF 04-JUN-1999; 99JP-0158571.
 XX
 PR 04-JUN-1999; 99JP-0158571.
 XX
 PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 XX
 DR WPI; 2001-260288/27.
 DR N-PSDB; AAF77276.
 XX
 PT A neutral amino acid transporter having broad substrate selectivity
 useful for in vitro analysis of permeation of drugs and poisons through
 cell membranes, blood and tissues -
 PT
 XX
 PS Claim 4; Page 17-18; 37pp; Japanese.
 XX
 CC This invention relates to a neutral amino acid transporter having broad
 substrate selectivity to neutral amino acids, and which is inactivated by
 an amino acid transport activating factor 4F2hc (4F2 heavy chain). The
 CC neutral amino acid transporter and its gene are useful in in vitro
 CC analysis of permeation of drugs and poisons through cell membrane, blood
 CC and tissues. The present sequence represents the rat L-type amino acid
 CC transporter 2 (LAT2).
 XX
 SQ Sequence 533 AA;
 Alignment Scores:

Pred. No.: 5.03e-116 Length: 533
 Score: 1108.00 Matches: 224
 Percent Similarity: 63.04% Conservative: 100
 Best Local Similarity: 43.58% Mismatches: 168
 Query Match: 26.89% Indels: 22
 DB: 22 Gaps: 5
 US-09-667-170A-440 (1-2239) x AAB72393 (1-533)
 QY 202 ATCTCAAGAGGTTTACCTGCGAGGAAATGTTAAACGGGAGGCTGCTTCCCTGGGCAAC 261
 Db 1 MetGluGlyGlyThrArgGlnArgAsnAsnThrAlaLysAsnHisProAspArgGlySer 20
 QY 262 AAGAGCCACCTGGGAGGAGCCGCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGACGCC 321
 Db 21 AspThrSerProGluAlaGluAlaSerSerGlyGlyGly----- 34
 QY 322 TTTCAGGAGAGAGAGAAAGTGCAGTGAAGAGAAAGTCACTTACTGAGGAGTCTCC 381
 Db 35 -----ValAlaLeuLysLysGluIleGlyLeuValSerAlaCysGly 48
 QY 382 ATTATCATTTGGCACCACCATTTGAGCAGCAATCTTCATCTCTCCTAAGGCGTCTCCAG 441
 Db 49 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 68
 QY 442 AACACGGGAGCGTGGGCGATGCTCTGACCATCTCGACGGTGTGGGGTCTCTGCTCACTA 501
 Db 69 AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyValIleThrAla 88
 QY 502 TTGGAGCTTTCTTATGCTGAATTTGGCAACACTATAAGAAATCTGGAGGTCACTAC 561
 Db 89 ValGlyAlaLeucystyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr 108
 QY 562 ACATATATTTTGAAGTCTTGTGTCATTTACGAGCTTTTGTACGAGTGGTGGCACTC 621
 Db 109 SerTyrValLysAspIlePheGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 128
 QY 622 CTCATAATACGCCCTGCGAGTACTCTCTGTATATCCCTGGCATTTGGAGCGTACATCTG 681
 Db 129 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 148
 QY 682 GAACCATTTTATTCAATGTGAATCCCTGAACTCCGATTCGATCAAGCTCATTACAGCTCG 741
 Db 149 GlnProLeuPheProThrCysPheProGluSerGlyLeuArgLeuLeuAlaIle 168
 QY 742 GCATAACTGTAGTATGTCCTAAATAGCATGATGTGTCAGCTGGAGCCCGGATCCAG 801
 Db 169 CysLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 188
 QY 802 ATTTCTTAACTTTTGAAGCTCACAGCAATCTGTGATAATATAGTCCCTGGAGTTATG 861
 Db 189 AspIlePheThrAlaGlyLysLeuLeuAlaLeuAlaIleIleMetGlyValVal 208
 QY 862 CAGCTAATTAAGGTCAACAGCAGAACTTT-----AAAGAGCCCTTTTCA 906
 Db 209 GlnIleCysLysGlyGlu-----PhePheTrpLeuGluProLysAsnAlaPheGlu 225
 QY 907 GGA---AGAGATTCAAGTATTACCGCGTGGCTGCTTTTATTTATGGAATGTATGCA 963
 Db 226 AsnPheGlnGluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAla 245
 QY 964 TATGCTGCTGTTTACCTCAACTTTGTTTACTGAAGAGTAGAAAAACCCCTGAAAAAAC 1023
 Db 246 TyrGlyGlyTrpAsnPheLeuAsnTyrValThrGluGluLeuValAspProTyrLysAsn 265
 QY 1024 ATTCCCTTGCATATGATATATCCATGGCCATTCACCAATTTGCTGCTGACAAAT 1083
 Db 266 LeuProArgAlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsn 285
 QY 1084 GTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAATGAGTGGGAGT 1143
 Db 286 IleAlaTyrValThrAlaMetSerProGlnGluLeuAlaSerAsnAlaValAlaVal 305

Db 28 GlySerGlyGlyGlyValAlaLeuLysLysGluIleGlyLeuValSerAlaCysGly 47
 QY 382 ATTATCATTTGGACCAATCATTTGGACGAGAAATCTTCATCTCCCTCCCAAGGGGTGCTCCAG 441
 Db 48 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 67
 QY 442 ACACGGGCGGTGGGCGATGCTCTGACCAATCTCGACGGTGTGGGGTCTGTCACCTA 501
 Db 68 AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyPheIleThrVal 87
 QY 502 TTGGAGCTTTTGTCTATCTGAATTTGGGAACTATATAAGAAATCTGGAGGCTATTAC 561
 Db 88 ValGlyAlaLeuCysThrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr 107
 QY 562 ACATATATTTTGAAGCTTTTGGTCCATTTACCACTTTTGTACGAGTCTGGGTGGAACCT 621
 Db 108 SerTyrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 127
 QY 622 CTCATATACGGCCCTGACGCTACTGCTGTGATATCCCTGGCATTTGGAGCTACATTCG 681
 Db 128 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 147
 QY 682 GAACCATTTTATTCATTAATGTAATCCCTGAACTTGGATCGATCAAGCTCATTTACAGCTGT 741
 Db 148 GlnProLeuPheProThrCysPheProGluSerGlyLeuArgLeuLeuAlaAlaIle 167
 QY 742 GGCATACATGTAGTGTCTTAATAGCATGCTGCTGAGTGGCGCCGGATCCAG 801
 Db 168 CysLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 187
 QY 802 ATTTCTTAACCTTTTGAAGCTACAGCAATTCATGTAATATAGTCCCTGGAGTATTG 861
 Db 188 AspIlePheThrAlaGlyLysLeuAlaLeuAlaLeuIleIleMetGlyIleVal 207
 QY 862 CAGCTAATTAAGGTCACAA-----ACGAGAACTTTAAGAGCGCTTTTCAGGA--AGA 912
 Db 208 GlnIleCysGlyGlyGlyTyrPheTrpLeuGluProLysAsnAlaPheGluAsnPheGln 227
 QY 913 GATTCAAGTATTACGGGGTGGCCACTGGCTTTTATATGGAATGATGATATGCTGGC 972
 Db 228 GluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAlaTyrGlyGly 247
 QY 973 TGGTTTACCTCAACTTTGTTACTCAAGAGTAGAAGACCTGAAACCAACCTCCCTT 1032
 Db 248 TrpAsnPheLeuAsnTyrValThrGluLeuValAlaAspProTyrLysAsnLeuProArg 267
 QY 1033 GCAATATGATATCCATGGCCATTCACCAATTTGGCTATGCTGCACAAATGTGGCCPAC 1092
 Db 268 AlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsnValAlaTyr 287
 QY 1093 TTTAGACCAATTAATGCTGAGAGCTGCTGCTTCAATGATGAGTGGCAGTGACCTTTCT 1152
 Db 288 ValThrAlaMetSerProGlnGluLeuAlaLeuAlaSerAsnAlaValAlaValThrPheGly 307
 QY 1153 GAGCGCTACTGGGAAATTTCTCATTAGCAGTTCGATCTTGTTCGCTCTCCCTGCTGCTT 1212
 Db 308 GluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeuSerThrPhe 327
 QY 1213 GCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTATTCTATGTTGGTCTGCTCGAG 1272
 Db 328 GlyGlyValAsnGlySerLeuPheThrSerSerArgLeuPhePheAlaGlyAlaArgGlu 347
 QY 1273 GGTACCTTCCAGAAATCTCTCCATGATGATGTCGCCAGCAGCAGCTCTCTACGAGCT 1332
 Db 348 GlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrProIleProAla 367
 QY 1333 GTTATTTGTTTGCACCCCTTTGACATGATATGCTCTCTCTGAGAGACCTCGACAGTCTT 1392
 Db 368 LeuLeuPheThrCysIleSerThrLeuLeuMetLeuValThrSerAspMetTyrThrLeu 387
 QY 1393 TTGAATTTCTCAGTTTGGCCAGGTGGCTTTTATGTTGGCTGGCAGTCTGCTGGCTGATT 1452
 Db 388 IleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAlaGlyGlnIle 407

QY 1453 TATCTTCATACAAATGCCAGATATGATCGTCTTCAAGGTGCACATGTTTCATCCCA 1512
 Db 408 ValLeuArgTrpLysLysProAspIleProArgProIleLysIleAsnLeuLeuPhePro 427
 QY 1513 GCTTTGTTTCCCTTCACATGCCTCTTCATGGTGGCCCTTTCCTCTATTCGGACCATTT 1572
 Db 428 IleIleTyrLeuLeuPheThrAlaPheLeuLeuValPheSerLeuTrpSerGluProVal 447
 QY 1573 AGTACAGGGATTGGCTTCGTTCATCACTCTGACCTGGAGTCCCTCGTATTCCTTTATT 1632
 Db 448 ValCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPheLeuGlyVal 467
 QY 1633 ATATGGCAAGAACCCAGGTGTTTAGAATAATGTCAGAGAAATAACACAGACATTA 1692
 Db 468 TyrTrpGlnHisLysProLysCysPheSerAspPheIleGluLeuThrLeuValSer 487
 QY 1693 CAATAATATCGAAGTTGTA---CCAGAAAGATAAG 1728
 Db 488 GlnLysMetCysValValValTyrProGluValGluArg 500
 RESULT 14
 AAY71061
 ID AAY71061 standard; Protein; 535 AA.
 XX
 AC AAY71061;
 XX
 DT 29-AUG-2000 (first entry)
 XX
 DE Human membrane transport protein, MTRP-6.
 XX
 KW Human; membrane transport protein; MTRP-6; antiinflammatory; cytostatic;
 KW antithyroid; immunosuppressive; thyromimetic; antidiabetic; nootropic;
 KW antidiarrheic; neuroprotective; antidepressant; nephrotropic; virucide;
 KW antihelminthic; protozoacide; antibacterial; neuroleptic; antigout;
 KW diagnosis; prevention; treatment; membrane transport disorder; epilepsy;
 KW Menkes disease; diabetes; Parkinson's disease; neurological disorder;
 KW Alzheimer's disease; depression; schizophrenia; immune disorder; allergy;
 KW inflammatory disorder; AIDS; Addison's disease; atherosclerosis; gout;
 KW Graves disease; Hashimoto's thyroiditis; microbial infection; cancer;
 KW cell proliferative disorder.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 11
 FT /note= "Phosphorylation site"
 FT Modified-site 61
 FT /note= "Phosphorylation site"
 FT Modified-site 103
 FT /note= "Phosphorylation site"
 FT Modified-site 109
 FT /note= "Phosphorylation site"
 FT Modified-site 179
 FT /note= "Phosphorylation site"
 FT Modified-site 232
 FT /note= "Phosphorylation site"
 FT Modified-site 305
 FT /note= "Phosphorylation site"
 FT Modified-site 337
 FT /note= "Phosphorylation site"
 FT Modified-site 363
 FT /note= "Phosphorylation site"
 FT Modified-site 487
 FT /note= "Phosphorylation site"
 FT Modified-site 502
 FT /note= "Phosphorylation site"
 FT Region 36..474
 FT /note= "Amino acid permeases signature sequence"
 FT Domain 76..96
 FT /label= Transmembrane_domain
 FT Domain 111..131
 FT /label= Transmembrane_domain

Db	28	GlySerGlyGlyGlyValAlaLeuLysGluIleGlyLeuValSerAlaCysGly	47
QY	382	ATTATCATTTGGACCACTCATTTGGACGAGAAATCTTCATCTCTCTAAGGCGTGCTCCAG	441
Db	48	IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu	67
QY	442	AACACGGGACGCTGGGATGCTCTGCACATCTGGACGGTGTGTGGGGTCTGTCACTA	501
Db	68	AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyPheIleThrVal	87
QY	502	TTTGGAGCTTTGCTCTATGCTGAATTTGGGAACAACATAAGAANAATCTGGAGGCTCATTC	561
Db	88	ValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr	107
QY	562	ACATATATTTGGAAAGTCTTTGGTCCATTACCACTTTTGTACGAGTCTGGGTGGAACTTC	621
Db	108	SerTyrValLysAspIlePheGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal	127
QY	622	CTCATTAATACGCCCTGCAGCTACTGCTGTGATATCCCTGCATTTGGACGCTACATCTG	681
Db	128	LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu	147
QY	682	GAACCATTTTTATTCAATGTGAATCCCTGAACTTCGATCAAGCTCAATACAGCTGTG	741
Db	148	GlnProLeuPheProThrCysPheProGluSerGlyLeuArgLeuLeuAlaIle	167
QY	742	GCATTAACCTAGTAGTGCTTAAATAGCATGAGTGCAGTGCAGGCGCCGGATCCAG	801
Db	168	CysLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln	187
QY	802	ATTTCTTTAACTTTTGCACAGCTCACAGCAATTTCTGATATTAATATAGTCCCTGGAGTTATG	861
Db	188	AspIlePheThrAlaGlyLysLeuLeuAlaLeuAlaLeuIleIleMetGlyIleVal	207
QY	862	CAGCTAATTAAGGTCAA-----ACGCAGAACTTTAAAGAGCGCTTTTTCAGGA--AGA	912
Db	208	GlnIleCysLysGlyGluTyrPheTrpLeuGluProLysAsnAlaPheGluAsnPheGln	227
QY	913	GATTCAAGTATTACCGGGTTGCCACTGGCTTTTATTATGGAAATGATGATGCTGGC	972
Db	228	GluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAlaTyrGlyGly	247
QY	973	TGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCCCTGAANAACCAATTCCTCTT	1032
Db	248	TrpAsnPheLeuAsnTyrValThrGluGluLeuValAspProTyrLysAsnLeuProArg	267
QY	1033	GCATATGTTATCCATGGCCATTGTCCACATTGGCTATGTGCTGCACAAATGTGCCTAC	1092
Db	268	AlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsnValAlaTyr	287
QY	1093	TTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAATGCGAGTGGCAGCTTTTCT	1152
Db	288	ValThrAlaMetSerProGlnGluLeuLeuAlaSerAsnAlaValAlaValThrPheGly	307
QY	1153	GAGCGGTACTTGGGAAATTTCTCATTAGCAGTTTCCGATCTTTGTTGGCTCTCTCTGCTT	1212
Db	308	GluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeuSerThrPhe	327
QY	1213	GGCTCCATGAACGGTGGTGGTTGCTGCTCCAGGTTATTCTATGTTGGCTCTCGAGAG	1272
Db	328	GlyGlyValAsnGlySerLeuPheThrSerSerArgLeuPhePheAlaGlyAlaArgGlu	347
QY	1273	GGTCACCTCCACAATCTCTCCATGATTCATGTGCCCAAGCACACTCCTCTACAGCT	1332
Db	348	GlyHisLeuProSerValLeuAlaMetIleHisValLysArgCysThrProIleProAla	367
QY	1333	GTTATTGTTTGTGACCCCTTTTGCAATGATTAATGCTTCTCTCTGGAGACTCGACAGCTT	1392
Db	368	LeuLeuPheThrCysIleSerThrLeuLeuMetLeuValThrSerAspMetTyrThrLeu	387
QY	1393	TTGAATTTCTCAGTTTCCAGGTTGGCTTTTATTGGCTGGCAGTTGTGGGCTGATT	1452
Db	388	IleAsnTyrValGlyPheIleAsnTyrLeuPheTyrGlyValThrValAlaGlyGlnIle	407

QY 1453 TATCTTCGATACAAATGCCAGATATGCATCGCTTTCAAGTGCCACTGTTCAATCCCA 1512
 Db 408 valLeuArgTrpLysProAspPheProIleLysIleAsnLeuLeuPhePro 427
 QY 1513 GCTTTGTTTCCATCGCTTCATCGGTTGCGCTTCCCTCTATTCGAGCCCAATT 1572
 Db 428 IleIleTyrLeuLeuPheTrpAlaPheLeuLeuValPheSerLeuTrpSerGluProVal 447
 QY 1573 AGTACAGGATTCGCTTCGTCATCTCTGACTGGAGTCCCTGCGTATTATCTTTATT 1632
 Db 448 valCysGlyIleGlyLeuAlaIleMetLeuThrGlyValProValTyrPheLeuGlyVal 467
 QY 1633 ATATGGCAGAAACCCAGGTGGTTAGTAATATGTCAGAGAAATAACACAGACATTA 1692
 Db 468 TyrTrpGlnHisLysProLysCysPheSerAspPheIleGluLeuLeuThrLeuValSer 487
 QY 1693 CAAATACTAGGAGTTGTA---CCAGAAGAGATAAG 1728
 Db 488 GlnLysMetCysValValValTyrProGluValGluArg 500

RESULT 15

AAB72395 standard; Protein; 535 AA.

ID AAB72395 standard; Protein; 535 AA.

XX AAB72395;

XX 24-MAY-2001 (first entry)

XX Human L-type amino acid transporter 2 (LAT2) amino acid sequence.

XX Neutral amino acid transporter; amino acid transport activating factor;

XX 4F2hc; 4F2 heavy chain; drug permeation; poison; human; LAT2;

XX L-type amino acid transporter 2.

XX Homo sapiens.

XX JP2000342270-A.

XX 12-DEC-2000.

XX 04-JUN-1999; 99JP-0158571.

XX 04-JUN-1999; 99JP-0158571.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI; 2001-260288/27.

XX N-PSDB; AAF77278.

XX A neutral amino acid transporter having broad substrate selectivity

XX useful for in vitro analysis of permeation of drugs and poisons through

XX cell membranes, blood and tissues -

XX Claim 4; Page 27-28; 37pp; Japanese.

XX This invention relates to a neutral amino acid transporter having broad

XX substrate selectivity to neutral amino acids, and which is inactivated by

XX an amino acid transport activating factor 4F2hc (4F2 heavy chain). The

XX neutral amino acid transporter and its gene are useful in vitro

XX analysis of permeation of drugs and poisons through cell membrane, blood

XX and tissues. The present sequence represents the human L-type amino acid

XX transporter 2 (LAT2).

XX Sequence 535 AA;

SQ

Alignment Scores:

Pred. No.:

Score: 2.76e-115 Length: 535

Matches: 220

Conservative: 103

Best Local Similarity: 44.62%

Mismatches: 161

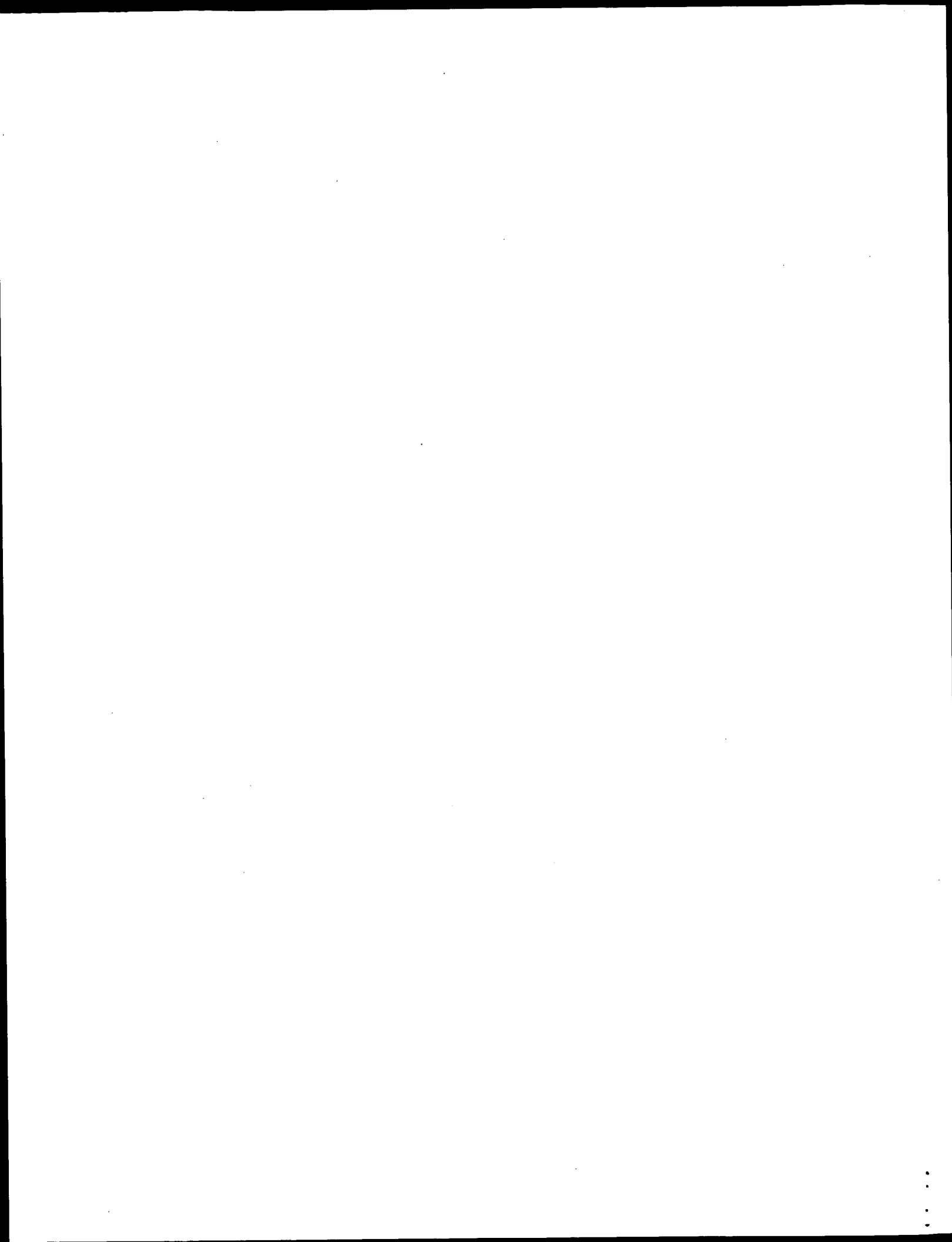
Indels: 9

DB: 22 Gaps: 4

US-09-667-170A-440 (1-2239) x AAB72395 (1-535)

QY 262 AAGAGCCACTGGGAGGAGCCCTTTTTCAGGAGAGACGCCCTTTTCAGGAGAGAGCC 321
 Db 13 LysLysHisProGly-----GlyGlyGluSerAspAlaSerProGluAla 27
 QY 322 TTTTCAGGAGAGAGAAAGTGCAGCTGAAGAGGAAAGTCTTCTACTGAGGGAGTCTCC 381
 Db 28 GlySerGlyGlyGlyValAlaLeuLysLysGluIleGlyLeuValSerAlaCysGly 47
 QY 382 ATTATCATTTGCCACCATTCATTGGAGCAGGAATCTTCTCTCTCTCTCTCTCTCTCT 441
 Db 48 IleIleValGlyAsnIleIleGlySerGlyIlePheValSerProLysGlyValLeuGlu 67
 QY 442 AACAGGGCCAGCTGGGCGATCTCTCTGACCATCTGAGGAGTGTGTGGTCTCTACTA 501
 Db 68 AsnAlaGlySerValGlyLeuAlaLeuIleValTrpIleValThrGlyPheIleThrVal 87
 QY 502 TTTGAGCTTTTCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCTATC 561
 Db 88 ValGlyAlaLeuCysTyrAlaGluLeuGlyValThrIleProLysSerGlyGlyAspTyr 107
 QY 562 ACATATATTTTGGAGTCTTTTGGTCCATTACAGCTTTTGTACGAGTCTGGTGGAACT 621
 Db 108 SerTyrValLysAspIlePheGlyGlyLeuAlaGlyPheLeuArgLeuTrpIleAlaVal 127
 QY 622 CTCATAATACGCCCTGCAGCTACTGCTGATATCCCTTGGCATTGGACGCTCATCTGT 681
 Db 128 LeuValIleTyrProThrAsnGlnAlaValIleAlaLeuThrPheSerAsnTyrValLeu 147
 QY 682 GAACCATTTTTCATCATGTAATCCTCAACTGGGATCGGATCAAGCTCATTACAGCTGTG 741
 Db 148 GlnProLeuPheProThrCysPheProGluSerGlyLeuArgLeuAlaIleAlaIle 167
 QY 742 GGCATAACTGTAGTGTCTCTAAATAGCATGAGTGTGAGTGTGAGCGCCGAGTCAAG 801
 Db 168 CysLeuLeuLeuLeuThrTrpValAsnCysSerSerValArgTrpAlaThrArgValGln 187
 QY 802 ATTTTCTTAACCTTTTGAAGCTCACAGCAATCTGTGAATATATAGTCCCTGGAGTTATG 861
 Db 188 AspIlePheThrAlaGlyLysLeuAlaLeuAlaLeuIleIleMetGlyIleVal 207
 QY 862 CAGCTAAATTAAGGTCAA-----ACGCAGAACTTTAAAGAGCGCTTTTCAGGA---AGA 912
 Db 208 GlnIleCysLysGlyGluTyrPheTrpLeuGluProLysAsnAlaPheGluAsnPheGln 227
 QY 913 GATTCAAGTATTACGCGGTTCGCTGCTTTTATTATGGAAATGATGATATGCTGGC 972
 Db 228 GluProAspIleGlyLeuValAlaLeuAlaPheLeuGlnGlySerPheAlaTyrGlyGly 247
 QY 973 TGGTTTACCTCAACTTTCTTACTGAAGAGTGAAGAACCCCTGAGAAACCACTCCCTT 1032
 Db 248 TrpAsnPheLeuAsnTyrValThrGluLeuValAspProTyrLysAsnLeuProArg 267
 QY 1033 GCATATGATATCCATGCCCATTTGTCACCATTTGCTGTGCTGATGTCACAAATGTGCGCTAC 1092
 Db 268 AlaIlePheIleSerIleProLeuValThrPheValTyrValPheAlaAsnValAlaTyr 287
 QY 1093 TTTTACGACCATTAATGCTGAGGAGTGTCTTCAATGTCAGTGGCAGTACCTTTCT 1152
 Db 288 ValThrAlaMetSerProGlnGluLeuLeuAlaSerAsnAlaValAlaValThrPheGly 307
 QY 1153 GAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCGATCTTGTGCGCTCTCTGCTTT 1212
 Db 308 GluLysLeuLeuGlyValMetAlaTrpIleMetProIleSerValAlaLeuSerThrPhe 327
 QY 1213 GGTCTCATGAACGCTGGTGTGTTTGTCTCCAGGTATTCTATGCTGCTCGCTCGAGAG 1272
 Db 328 GlyClyValAsnGlySerLeuPheThrSerSerArgLeuPhePheAlaGlyAlaArgGlu 347
 QY 1273 GGTCACTTCCAGAAATTCCTCTCCATGATTCATGTCGCGAAGCACACTCTCTTACAGCT 1332

Search completed: April 16, 2003, 16:09:32
Job time : 86 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 09:34:04 ; Search time 6504 Seconds
(without alignments)
10018.632 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table: OLIGO_NUC

Gapop 60.0 , Capext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Result No.	Score	Query Match	Length	DB ID	Description
1	2239	100.0	2239	6	AX321909 Sequence
2	1867	83.4	5981	6	AX321910 Sequence
3	1816	81.1	3144	9	HS2277882
4	1722	76.9	2482	9	AF252872 Homo sapi
5	1704	76.1	2155	9	BC012087 Homo sapi
6	1459	65.2	1874	9	AF200708 Homo sapi
7	1454	64.9	1861	9	AB026891 Homo sapi
8	1339	59.8	2000	9	AB040875 Homo sapi
9	479	21.4	64196	9	AC110804 Homo sapi
10	281	12.6	337	6	AX321911 Sequence
11	269	12.0	596	6	AX351341 Sequence
12	182	8.1	161280	9	AC093903 Homo sapi
13	173	7.7	1683	9	AB042201 Homo sapi
14	128	5.7	3994	9	AC116610 Homo sapi
15	107	4.8	575	6	AX341060 Sequence
16	84	3.8	632	11	G61322 SHGC-85475
17	52	2.3	492	10	AB03765504 Mus muscu
18	52	2.3	2216	10	AB022345 Mus muscu
19	52	2.3	183949	2	AC118181 Rattus no
20	49	2.2	724	9	HSNC2ALPH
21	49	2.2	788	9	BC010025 Homo sapi
22	45	2.0	983	6	AX011754 Sequence
23	45	2.0	92203	2	AF001920 Homo sapi
24	45	2.0	162511	9	AC009470 Homo sapi
25	45	2.0	182547	9	AF001201 Homo sapi
26	39	1.7	942	10	AB03765503 Mus muscu
27	37	1.7	97773	2	AC106406 Rattus no
28	37	1.7	183949	2	AC118181 Rattus no
29	35	1.6	40690	9	AC112186 Homo sapi
30	35	1.6	70638	9	AC079306 Homo sapi
31	35	1.6	110000	2	AC023632_0 Homo sapi
32	35	1.6	110000	2	AC023632_2 Homo sapi
33	35	1.6	143183	2	AC026927 Continuation (3 of
34	35	1.6	145695	2	AC026927 Homo sapi
35	35	1.6	159692	2	AC013693 Homo sapi
36	35	1.6	161806	2	AC016495 Homo sapi
37	35	1.6	195623	9	AF004218 Homo sapi
38	34	1.5	149336	2	AF315312 Homo sapi
39	34	1.5	159104	9	AC108925 Homo sapi
40	33	1.5	33667	9	HSU227D1 Homo sapi
41	33	1.5	38041	9	AC002985 Human DNA s
42	33	1.5	39617	9	AF000348 Homo sapi
43	33	1.5	43472	2	AC066605 Homo sapi
44	33	1.5	44430	9	AC092309 Homo sapi
45	33	1.5	44869	9	AB022537 Homo sapi

ALIGNMENTS

RESULT 1	AX321909	Sequence 440 from Patent WO0172295.	linear	PAT 15-DEC-2001
LOCUS	AX321909	GI:17906515		
DEFINITION	Sequence 440 from Patent WO0172295.			
ACCESSION	AX321909			
VERSION	AX321909.1			
KEYWORDS	human.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1			
AUTHORS	Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,			
	Indrias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,			
	Mannion,J. and Kalos,M.D.			

Pred. No. is the number of results predicted by chance to have a

TITLE	Compositions and methods for the therapy and diagnosis of lung cancer									
JOURNAL	Patent: WO 017295-A 440 04-OCT-2001;									
FEATURES	CORIXA CORPORATION (US)									
source	Location/Qualifiers									
	1. .2239									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
BASE COUNT	619 a	444 c	493 g	683 t						
ORIGIN										
Query Match	100.0%;	Score	2239;	DB	6;	Length	2239;			
Best Local Similarity	100.0%;	Pred.	No. 0;							
Matches 2239;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;		
QY	1	GGAGGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAC	60							
Db	1	GGAGGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAC	60							
QY	61	AGAAATTAAGAAAAAAGAAAGAAAGAAAGAGAGAGAGAAATCCAGGCCAATTTGTGG	120							
Db	61	AGAAATTAAGAAAAAAGAAAGAAAGAAAGAGAGAGAGAAATCCAGGCCAATTTGTGG	120							
QY	121	CATAGATTTTATCATATCTCTGGATTTTTCGATTTCTTTTTCATCATCTGGATTCA	180							
Db	121	CATAGATTTTATCATATCTCTGGATTTTTCGATTTCTTTTTCATCATCTGGATTCA	180							
QY	181	GGAAAGCCTGTGTGTCACCATCTCCAAAGAGAGTTTACCTGCAGGGAATGTTAACGGG	240							
Db	181	GGAAAGCCTGTGTGTCACCATCTCCAAAGAGAGTTTACCTGCAGGGAATGTTAACGGG	240							
QY	241	AGGCTGCTTCCCTGGGCAACAGGAGGCCACTGGGGAGGAGCGCTTTTCAGGAAGAGAC	300							
Db	241	AGGCTGCTTCCCTGGGCAACAGGAGGCCACTGGGGAGGAGCGCTTTTCAGGAAGAGAC	300							
QY	301	GCCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGACGCTGAAGAGAAAGTC	360							
Db	301	GCCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGACGCTGAAGAGAAAGTC	360							
QY	361	ACTTTACTGAGGGAGTCTCCATATCATTTGGACCATCATTTGGAGCAGGAATCTTCATC	420							
Db	361	ACTTTACTGAGGGAGTCTCCATATCATTTGGACCATCATTTGGAGCAGGAATCTTCATC	420							
QY	421	TCTCCTAAGGCGTCTCCAGAACACGGGAGCGTGGGCATGTCTCTGACCATCTGGACG	480							
Db	421	TCTCCTAAGGCGTCTCCAGAACACGGGAGCGTGGGCATGTCTCTGACCATCTGGACG	480							
QY	481	GTGTGTGGGTCCTGTCACTATTTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATA	540							
Db	481	GTGTGTGGGTCCTGTCACTATTTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATA	540							
QY	541	AAGAAATCTGGAGTCTATTACACATATATTTTGAAGTCTTTGGTCCATTACCACTTTT	600							
Db	541	AAGAAATCTGGAGTCTATTACACATATATTTTGAAGTCTTTGGTCCATTACCACTTTT	600							
QY	601	GTACGAGTCTGGGTGGAACTCCTCATATACGCCCTGCAGCTACTGCTGTGATATCCCTG	660							
Db	601	GTACGAGTCTGGGTGGAACTCCTCATATACGCCCTGCAGCTACTGCTGTGATATCCCTG	660							
QY	661	GCATTGGACGCTACATTTCTGGACCATTTTATTCATGTAATTTCCCTGAACTTTCGG	720							
Db	661	GCATTGGACGCTACATTTCTGGACCATTTTATTCATGTAATTTCCCTGAACTTTCGG	720							
QY	721	ATCAAGCTCATTTACAGCTGTGGCAATACTGTAGTGTGGTCCCTAAATAGCATGATGTC	780							
Db	721	ATCAAGCTCATTTACAGCTGTGGCAATACTGTAGTGTGGTCCCTAAATAGCATGATGTC	780							
QY	781	AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTCCAGCTCACAGCAATTTCTGATA	840							
Db	781	AGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTCCAGCTCACAGCAATTTCTGATA	840							
QY	841	ATTATAGTCCCTGGAGTTATGACGCTAATTAAGGTCAAAGCAGAACTTTAAAGAGCC	900							

Db	841	ATTATAGTCCCTGGAGTTATGACGCTAATTAAGGTCAACAGCAAACTTTAAGAGCCC	900
Qy	901	TTTTCAGGAAGAGATTCAAGTATTACCGGTTGCCACTGGCTTTTATTATGAATGTAT	960
Db	901	TTTTCAGGAAGAGATTCAAGTATTACCGGTTGCCACTGGCTTTTATTATGAATGTAT	960
Qy	961	GCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCCCTGAAAA	1020
Db	961	GCATATGCTGGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCCCTGAAAA	1020
Qy	1021	ACCATTCCTTCCCAATATGATATCCATGGCCATTTGCACATTGGCTATGTCTGACA	1080
Db	1021	ACCATTCCTTCCCAATATGATATCCATGGCCATTTGCACATTGGCTATGTCTGACA	1080
Qy	1081	AATGGGCTTACTTTACGACCAATTAATGCTGAGGAGTGTCTTCAATATGCAAGTGACA	1140
Db	1081	AATGGGCTTACTTTACGACCAATTAATGCTGAGGAGTGTCTTCAATATGCAAGTGACA	1140
Qy	1141	GTGACCTTTCTGAGCGGCTACTGGGAAATTTCTCATATGACAGTTCCGATCTTTTGGCC	1200
Db	1141	GTGACCTTTCTGAGCGGCTACTGGGAAATTTCTCATATGACAGTTCCGATCTTTTGGCC	1200
Qy	1201	CTCTCTCTGTTGGCTCCATGAACGGTGTGTGTTGCTGTCTCCAGGTTATTTCTATGTT	1260
Db	1201	CTCTCTCTGTTGGCTCCATGAACGGTGTGTGTTGCTGTCTCCAGGTTATTTCTATGTT	1260
Qy	1261	GGCTCTCGAGAGGCTACCTTTCCAGAAATCTCTCCATGATTCATGTCGCGCAAGCACAT	1320
Db	1261	GGCTCTCGAGAGGCTACCTTTCCAGAAATCTCTCCATGATTCATGTCGCGCAAGCACAT	1320
Qy	1321	CCCTACAGCTGTTTATTTGTTTGGACCTTTTGACAAATGATAATGCTCTCTCTGGAGAC	1380
Db	1321	CCCTACAGCTGTTTATTTGTTTGGACCTTTTGACAAATGATAATGCTCTCTCTGGAGAC	1380
Qy	1381	CTCGACAGCTTTTCAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGCTGGCAGTT	1440
Db	1381	CTCGACAGCTTTTCAATTTCTCAGTTTGGCAGGTGGCTTTTATTGGCTGGCAGTT	1440
Qy	1441	GCTGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGCTCTTCAAGGTGCA	1500
Db	1441	GCTGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGCTCTTCAAGGTGCA	1500
Qy	1501	CTGTTCATCCAGCTTTGTTTCTTCATGCTTCTCATGTTGGCTTTCCCTCTAT	1560
Db	1501	CTGTTCATCCAGCTTTGTTTCTTCATGCTTCTCATGTTGGCTTTCCCTCTAT	1560
Qy	1561	TCGGACCAATTTAGTACAGGATTTGGCTTCATCTCTGCTGAGTCCCTCGGTAT	1620
Db	1561	TCGGACCAATTTAGTACAGGATTTGGCTTCATCTCTGCTGAGTCCCTCGGTAT	1620
Qy	1621	TATCTCTTTTATATATGAGCAAGAACCCAGTGGTTTAGAATAATGTCAGAGAAAAA	1680
Db	1621	TATCTCTTTTATATATGAGCAAGAACCCAGTGGTTTAGAATAATGTCAGAGAAAAA	1680
Qy	1681	ACCAGAACATTAACAAATAACTTGAAGTTGTACCAAGAAAGATAAGTATGAATAAT	1740
Db	1681	ACCAGAACATTAACAAATAACTTGAAGTTGTACCAAGAAAGATAAGTATGAATAAT	1740
Qy	1741	GGACTTTGAGATCTTGCAATCTGCCCAAGGGAGACAAAAATAGGATTTTACTTCAT	1800
Db	1741	GGACTTTGAGATCTTGCAATCTGCCCAAGGGAGACAAAAATAGGATTTTACTTCAT	1800
Qy	1801	TTTTCTCAAGCTCTAGAGAAATTAACAATTTGGTGAATAACAAAGAGTCAAGTATTTTA	1860
Db	1801	TTTTCTCAAGCTCTAGAGAAATTAACAATTTGGTGAATAACAAAGAGTCAAGTATTTTA	1860
Qy	1861	TTCATATATTTTACGATATTCGAACTAATTTCTTAAGAAATTTAGTTATAACTCTATGAT	1920
Db	1861	TTCATATATTTTACGATATTCGAACTAATTTCTTAAGAAATTTAGTTATAACTCTATGAT	1920
Qy	1921	TTATAGAAAGTGAATATGACGTTATTTCTATGAGTCCCAAAATTTCTGAGTCTCTGATACC	1980
Db	1921	TTATAGAAAGTGAATATGACGTTATTTCTATGAGTCCCAAAATTTCTGAGTCTCTGATACC	1980

QY 1174 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCTCTGCTCCATGAACGGTGGTGTG 1233
 Db 946 TCATTAGCAGTCCGATCTTTGTTGCCCTCTCTCTGCTCCATGAACGGTGGTGTG 1005
 QY 1234 TTTGCTGCTCCAGGTATTCTATGTTGGCTCGAGAGGTCACCTTCCAGAAATCCTC 1293
 Db 1006 TTTGCTGCTCCAGGTATTCTATGTTGGCTCGAGAGGTCACCTTCCAGAAATCCTC 1065
 QY 1294 TCCATGATTATGTCGGAAGCACACTCTCTACACAGCTGTTATTGTTTTCGACCCCTTG 1353
 Db 1066 TCCATGATTATGTCGGAAGCACACTCTCTACACAGCTGTTATTGTTTTCGACCCCTTG 1125
 QY 1354 ACAATGATAATGCTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1413
 Db 1126 ACAATGATAATGCTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1185
 QY 1414 AGTGGCTTTTATGAGGCTGCGAGTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1473
 Db 1186 AGTGGCTTTTATGAGGCTGCGAGTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1245
 QY 1474 GATATGATCGCTCTTCAAGTGGCAGTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1533
 Db 1246 GATATGATCGCTCTTCAAGTGGCAGTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1305
 QY 1534 CTCCTCATGTTGCTCTTCAAGTGGCAGTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1593
 Db 1306 CTCCTCATGTTGCTCTTCAAGTGGCAGTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1365
 QY 1594 ATCACTCTGAGTGGCTGCGAGTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1653
 Db 1366 ATCACTCTGAGTGGCTGCGAGTCTCTGAGACCTCGACAGTCTTTTGAATTCCTCAGTTTGGC 1425
 QY 1654 TGGTTAGATAATGTCGAGAGAAATACCAAGACATTTACAAATATCTGGAAGTTGTA 1713
 Db 1426 TGGTTAGATAATGTCGAGAGAAATACCAAGACATTTACAAATATCTGGAAGTTGTA 1485
 QY 1714 CCAGAGAGATAGTATGAACTTAAGTGGCTGAGATCTTGGCAATCTGCCCAAGGGA 1773
 Db 1486 CCAGAGAGATAGTATGAACTTAAGTGGCTGAGATCTTGGCAATCTGCCCAAGGGA 1545
 QY 1774 GACAAATAGGATTTTACTTCAATTTCTGAAAGTCTAGAGAAATACAACTTTGGTG 1833
 Db 1546 GACAAATAGGATTTTACTTCAATTTCTGAAAGTCTAGAGAAATACAACTTTGGTG 1605
 QY 1834 ATAAACAAAGGAGTCACTTATTTTATTCATATATTTTATGCAATATTCGAATTTCT 1893
 Db 1606 ATAAACAAAGGAGTCACTTATTTTATTCATATATTTTATGCAATATTCGAATTTCT 1665
 QY 1894 AAGAAATTTAGTATTAATCTATGATGATATGAAAGTGAATATGCAATTTATCTATGAG 1953
 Db 1666 AAGAAATTTAGTATTAATCTATGATGATATGAAAGTGAATATGCAATTTATCTATGAG 1725
 QY 1954 TCGCACAATTTCTGAGTCTCTGATACCTACTATTGGGGTTAGAGAAAGAGCTAGACAA 2013
 Db 1726 TCGCACAATTTCTGAGTCTCTGATACCTACTATTGGGGTTAGAGAAAGAGCTAGACAA 1785
 QY 2014 TTACTATGTTGATCTCTACACATATGTTAGCAGCGCAAGAACCTTCAATTTGAAG 2073
 Db 1786 TTACTATGTTGATCTCTACACATATGTTAGCAGCGCAAGAACCTTCAATTTGAAG 1845
 QY 2074 ACTGAGATTTTCTGATATATGTTTCTGAAAGTGGTTTACACACTACAGATGTCT 2133
 Db 1846 ACTGAGATTTTCTGATATATGTTTCTGAAAGTGGTTTACACACTACAGATGTCT 1905
 QY 2134 ATACTGGAAGTGGTTTCAATTTCTGAAAGGAGTACATCATGATTTATGCAAGAG 2193
 Db 1906 ATACTGGAAGTGGTTTCAATTTCTGAAAGGAGTACATCATGATTTATGCAAGAG 1965
 QY 2194 GAGAGAA 2200
 Db 1966 GAGAGAA 1972

RESULT 4
 AF252872
 LOCUS
 DEFINITION Homo sapiens cystine/glutamate transporter xct mRNA linear PRI 02-MAY-2001
 AF252872
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 2482)
 AUTHORS Chancy,C.D., Kekuda,R., Wang,H., Huang,W., Prasad,P.D., Smith,S.B.
 and Ganapathy,V.
 TITLE Structure, Function and Regulation of Human Cystine/Glutamate
 Transporter in Retinal Pigment Epithelial Cells
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2482)
 AUTHORS Wang,H., Prasad,P.D. and Ganapathy,V.
 TITLE Direct Submission
 JOURNAL Submitted (05-APR-2000) Biochemistry & Molecular Biology, Medical
 College of Georgia, 1120 15th Street, Augusta, GA 30912, USA
 FEATURES
 Location/Qualifiers
 1..2482
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="placenta"
 232..1737
 /codon_start=1
 /product="cystine/glutamate transporter xct"
 /protein_id="AAK4911.1"
 /db_xref="GI:13924720"
 /translation="MVRPVSSTISKGYLOGNVRLPSLGNKEPPOEOKVQLRRKY
 TLRKGSIIIGTIGAGIFIPKQLQNTSVGMSLTITWVGVLSTLFGALSAYELGT
 TLKSGSHYTYLEVGPLPAFVRYVVELLIIRPAATVLSLAFGRYLEPFFIQCEI
 PELAIKLITAVGTYVMVLSMSVSWSARIQITFLFKLITALLIIVPGVMOLIKGT
 QNFKDFSGROSSITRPLAFYVMYAVAGWFLNFVEENPEKTIPLAICISMAI
 VTIGVLTNVAFTTINAEILLLSNAVATPSERLNGFNLSLAVPIFVALSCFGSNGG
 VFVSRIFYVASREGLPELISLHVRKHTPLPVIIVLHPTMLFSGDLDLNLNPL
 SFARMLFTGLAVAGLITRYKCPDMRPFKVPFLFIPALFSTCLFMVALSYSDPFT
 GIGFVITLTGVPAYYLFITWDKKRPFIMSEKTRITLQIILEVVPEDKL"
 735 a 484 c 521 g 742 t
 BASE COUNT
 ORIGIN
 Query Match 76.9%; Score 1722; DB 9; Length 2482;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1772; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 334 GAGAAAGTCAGCTGAAGAGGAAAGTCATTTACTGAGGGAGTCTCCATTTATCATTTGGC 393
 Db 337 GAGAAAGTCAGCTGAAGAGGAAAGTCATTTACTGAGGGAGTCTCCATTTATCATTTGGC 396
 QY 394 ACCATCATTTGGAGCAGGAATCTTCTCTCTTAAGGGCGTCTCCAGAACACGGGCAGC 453
 Db 397 ACCATCATTTGGAGCAGGAATCTTCTCTCTTAAGGGCGTCTCCAGAACACGGGCAGC 456
 QY 454 GTGGGCATCTCTGACCATCTGGAGGCTGTGTGGGCTCTCTGACACTATTTGGAGCTTTG 513
 Db 457 GTGGGCATCTCTGACCATCTGGAGGCTGTGTGGGCTCTCTGACACTATTTGGAGCTTTG 516
 QY 514 TCTTATGCTGAATTTGGGAACAACACTATAAGAAATCTGGAGGTCATTACACATATATTTG 573
 Db 517 TCTTATGCTGAATTTGGGAACAACACTATAAGAAATCTGGAGGTCATTACACATATATTTG 576
 QY 574 GAAGCTTTGGTCCCATTTACCAGCTTTTGTACAGTCTGGTGGAACTCGCTCATATATAGC 633
 Db 577 GAAGCTTTGGTCCCATTTACCAGCTTTTGTACAGTCTGGTGGAACTCGCTCATATATAGC 636
 QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 693
 Db 637 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTTT 696
 QY 694 ATTCAATGTGAATCCCTGAACCTTGGCATTCAGCTCATTACAGCTGTGGGCATATAGTGA 753

Db	697	ATTCAATGGAATCCCTGAACTTGGCATCAAGCTCATTTACAGCTGTGGGCAATAACTGTA	756		Db	1777	GACACAAATAGGGATTTTACTTCATTTTCTGAAAGTCTAGAGAAATACAACTTTGGTG	1836
Qy	754	GTGATGGTCTTAATAGCATGAGTGTGAGCTGAGGCGCCGGATCCAGATTTTCTTAACC	813		Qy	1834	ATAACAAAAGGAGTCACTTATTTTATTCATATATTTTAGCATATTCGAACATAATTTCT	1893
Db	757	GTGATGGTCTTAATAGCATGAGTGTGAGCTGAGGCGCCGGATCCAGATTTTCTTAACC	816		Db	1837	ATAACAAAAGGAGTCACTTATTTTATTCATATATTTTAGCATATTCGAACATAATTTCT	1896
Qy	814	TTTTGCAAGCTCACAGCAATTTCTGATAATATATAGTCCCTGGAGTTATGACAGCTAAATTA	873		Qy	1894	AAGAAATTTAGTTATAACTCTATGATAGTATAGAAAGTGAATATGCGATTTATCTATGAG	1953
Db	817	TTTTGCAAGCTCACAGCAATTTCTGATAATATATAGTCCCTGGAGTTATGACAGCTAAATTA	876		Db	1897	AAGAAATTTAGTTATAACTCTATGATAGTATAGAAAGTGAATATGCGATTTATCTATGAG	1956
Qy	874	GGTCAAGCGAAGACTTTAAAGAGCCCTTTTCAGGAAGAGATTCAGATTAATGCGGGTTG	933		Qy	1954	TCGCACAAATTTGAGTCTCTCATACCTACCTATTTGGGTTAGGAGAAAAGACTAGACAA	2013
Db	877	GGTCAAGCGAAGACTTTAAAGAGCCCTTTTCAGGAAGAGATTCAGATTAATGCGGGTTG	936		Db	1957	TCGCACAAATTTGAGTCTCTCATACCTACCTATTTGGGTTAGGAGAAAAGACTAGACAA	2016
Qy	934	CCACTGGCTTTTATATGGAATCTATGATATGCTGCGCTGGTTTACCTCAACTTTGTT	993		Qy	2014	TTACTATGTGTCATTTCTCTCAACATATGTTAGCACGCAAGAACCTTCAAAATTTGAAG	2073
Db	937	CCACTGGCTTTTATATGGAATCTATGATATGCTGCGCTGGTTTACCTCAACTTTGTT	996		Db	2017	TTACTATGTGTCATTTCTCTCAACATATGTTAGCACGCAAGAACCTTCAAAATTTGAAG	2076
Qy	994	ACTGAAGAGTAGAAAACCCCTGAAAAACCATTCCTTGCATATATGATATCATATGCGC	1053		Qy	2074	ACTGAGATTTTCTGTATATATGCGGTTTGTAA	2106
Db	997	ACTGAAGAGTAGAAAACCCCTGAAAAACCATTCCTTGCATATATGATATCATATGCGC	1056		Db	2077	ACTGAGATTTTCTGTATATATGCGGTTTGTAA	2109
Qy	1054	ATTGTACCAATGGTATGTGCTGACAAATGTGGCCCTACTTTAGGACCAATTAATGCTGAG	1113		RESULT 5			
Db	1057	ATTGTACCAATGGTATGTGCTGACAAATGTGGCCCTACTTTAGGACCAATTAATGCTGAG	1116		BC012087			
Qy	1114	GAGTGTGCTTTCAATAGCATGCGAGTGCCTTTTCTGAGCGGCTACCTGCGAAATTTTC	1173		LOCUS			
Db	1117	GAGTGTGCTTTCAATAGCATGCGAGTGCCTTTTCTGAGCGGCTACCTGCGAAATTTTC	1176		DEFINITION			
Qy	1174	TCATTAGCAGTTCGGATCTTTTGGCCCTCTCTGCTTTGGCTCCATGAACGGTGGTG	1233		ACCESSION			
Db	1177	TCATTAGCAGTTCGGATCTTTTGGCCCTCTCTGCTTTGGCTCCATGAACGGTGGTG	1236		BC012087			
Qy	1234	TTTGTGTCTCAGGTTATCTATGTTGCGTCTCGAGAGGTCACTTCAGAAATCCTC	1293		VERSION			
Db	1237	TTTGTGTCTCAGGTTATCTATGTTGCGTCTCGAGAGGTCACTTCAGAAATCCTC	1296		KEYWORDS			
Qy	1294	TCCATGATTCATGTCGCAAGCACACTCTCTACGAGCTGTATGTTTGGACCCCTTG	1353		SOURCE			
Db	1297	TCCATGATTCATGTCGCAAGCACACTCTCTACGAGCTGTATGTTTGGACCCCTTG	1356		ORGANISM			
Qy	1354	ACAATGATAATGCTCTCTCGAGACCTCGACAGCTCTTTTGAATTTCCCTCAGTTTGGC	1413		REFERENCE			
Db	1357	ACAATGATAATGCTCTCTCGAGACCTCGACAGCTCTTTTGAATTTCCCTCAGTTTGGC	1416		AUTHORS			
Qy	1414	AGGTGGCTTTTATGCGCTGCGAGTGTGCTGCGCTGATTTATCTTCGATACAAATGCCA	1473		TITLE			
Db	1417	AGGTGGCTTTTATGCGCTGCGAGTGTGCTGCGCTGATTTATCTTCGATACAAATGCCA	1476		JOURNAL			
Qy	1474	GATATGCATCGCTCTTCAAGGTGCACTGTTCATCCAGCTTTGTTTCTTCACATGC	1533		REMARK			
Db	1477	GATATGCATCGCTCTTCAAGGTGCACTGTTCATCCAGCTTTGTTTCTTCACATGC	1536		COMMENT			
Qy	1534	CTCTTCATGCTGCTTCCCTCTATTCGAGCCATTTAGTACAGGATTTGGCTTGGT	1593					
Db	1537	CTCTTCATGCTGCTTCCCTCTATTCGAGCCATTTAGTACAGGATTTGGCTTGGT	1596					
Qy	1594	ATCACTCTGAGTGGAGTCCCTGCGTATTTATCTTTATATATGAGCAAGAAACCCAGG	1653					
Db	1597	ATCACTCTGAGTGGAGTCCCTGCGTATTTATCTTTATATATGAGCAAGAAACCCAGG	1656					
Qy	1654	TGTTTTAGAAATATGTCAGAGAAAATAACCAAGCAATTTACAAATAATCTGGAAGTTGTA	1713					
Db	1657	TGTTTTAGAAATATGTCAGAGAAAATAACCAAGCAATTTACAAATAATCTGGAAGTTGTA	1716					
Qy	1714	CCAGAGAGAGATAAGTTATGAATTAATGAGCTTGAGATCTTGGCAATTCGCCAAGGGGA	1773					
Db	1717	CCAGAGAGAGATAAGTTATGAATTAATGAGCTTGAGATCTTGGCAATTCGCCAAGGGGA	1776					
Qy	1774	GACACAAATAGGATTTTACTTCATTTCTGAAAGTCTAGAGAAATACAACTTTGGTG	1833					

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 29 Row: 9 Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5668544.

FEATURES

Source

Location/Qualifiers
1..2155
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:20026 IMAGE:4562994"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH_MGC_14"

/lab_host="DH10B-R"		Db		986																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
---------------------	--	----	--	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 2066 ATACTGTGAAAAGTGTGTTTCAATCTCTGAAAAAGCATACATCATGATTATGGCAA 2122

RESULT 6
AF200708
LOCUS
DEFINITION Homo sapiens calcium channel blocker resistance protein CCBRI mRNA, complete cds.
ACCESSION AF200708
VERSION AF200708.1 GI:11493651
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1874)
Konkin, D.S. and Beach, D.H.
CCBR1, novel CD98 light chain implicated in redox control and calcium signaling
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 1874)
Konkin, D.S. and Beach, D.H.
Direct Submission
TITLE
JOURNAL
FEATURES
source
1. .1874
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="fibrosarcoma HT1080"
246. 1751
/note="X-c transporter; CD98 light chain"
/codon_start=1
/product="calcium channel blocker resistance protein CCBRI"
/protein_id="AAG35592.1"
/db_xref="GI:11493652"
/translations="MVRKPVSTISKGYLOGNVNRLPSLGNKPPGQEVOLRRKV
TLRGVLIIGTIGAGIFLSPKVLQNTGSVGMSLTINTVCGVLSLFGALSYARELT
TKRSGHYTHILEVPGPLPAFVRVVELLIIRPATATVISLAFGRYILEPFFIOCEI
PELAILTAVGTIVVVLNLSVSWARIQIFLTPCKTALILIIIPVGMQIIRKGT
ONKDFADROSSITILPLAFYGYMYAGWFLNFVEENPEKTIPLAICISMAI
VTGYVLTVAAYETHINABELLSNAYATFBERLGNFSLAVPIFVALSCRSXNGG
VFASVLYVNSREGLPEILSMIHVRKHPTLPVAVLHPLTMILFSGDLSDLLNFI
SFARWFLGLAVGLIYLRKCPDMRPFKPLFIPLAFSEFCLEWVALSLYSDFEST
GIGFVITLGVPAAYLFIILWDKPRFRIMSEKITFTLIQLILEVVPEDKL"
BASE COUNT 490 a 407 c 422 g 555 t

Query Match 65.2%; Score 1459; DB 9; Length 1874;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATGGC 393
DB 351 GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATGGC 410

QY 394 ACCATCATTTGAGCAGGAATCTTCATCTCTAAGCGGTGCTCCAGACAGCGGCAGC 453
DB 411 ACCATCATTTGAGCAGGAATCTTCATCTCTAAGCGGTGCTCCAGACAGCGGCAGC 470

QY 454 GTGGCATGTCTCTGACCATCTGACCGGTGTGTGGGGTCTCTGTCACATATTGGAGCTTTG 513
DB 471 GTGGCATGTCTCTGACCATCTGACCGGTGTGTGGGGTCTCTGTCACATATTGGAGCTTTG 530

QY 514 TCTTATGCTGAATTTGGGAACAACTATTAAGAAATCTGGAGTCAATACATATATTG 573
DB 531 TCTTATGCTGAATTTGGGAACAACTATTAAGAAATCTGGAGTCAATACATATATTG 590

QY 574 GAAGTCTTTGTTGCTTACAGCTTTTCTACAGCTCTGGGTGGAACTCTCATATAATACGC 633
DB 591 GAAGTCTTTGTTGCTTACAGCTTTTCTACAGCTCTGGGTGGAACTCTCATATAATACGC 650

QY 634 CCGTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATCTCTGGAACCAATTTTT 693
DB 651 CCGTGCAGCTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATCTCTGGAACCAATTTTT 710

QY 694 ATTCATGTGAAATCCCTGGAATCTGCGATCAAGCTCATTACAGCTGTGGGCATAACTGTA 753
DB 711 ATTCATGTGAAATCCCTGGAATCTGCGATCAAGCTCATTACAGCTGTGGGCATAACTGTA 770

QY 754 GTGATGGTCTAAATAGCATGAGTGTGAGCGCCCGGATCCAGATTTTCTTAACC 813
DB 771 GTGATGGTCTAAATAGCATGAGTGTGAGCGCCCGGATCCAGATTTTCTTAACC 830

QY 814 TTTTCAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTATGACGCTAATATAA 873
DB 831 TTTTCAAGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTATGACGCTAATATAA 890

QY 874 GGTCAAGCGCAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAGATTTACCGGTTG 933
DB 891 GGTCAAGCGCAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAGATTTACCGGTTG 950

QY 934 CCACCTGGCTTTTATTATGGAATGTATCATATGCTGGCTGGTTTACCTCAACTTTGTT 993
DB 951 CCACCTGGCTTTTATTATGGAATGTATCATATGCTGGCTGGTTTACCTCAACTTTGTT 1010

QY 994 ACTGAAGAGTAGAAAACCCCTGAAAAACCAATTCCTCTGCAATATGATATCATGGCC 1053
DB 1011 ACTGAAGAGTAGAAAACCCCTGAAAAACCAATTCCTCTGCAATATGATATCATGGCC 1070

QY 1054 ATTGTCAACATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1113
DB 1071 ATTGTCAACATTTGGCTATGTGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1130

QY 1114 GAGCTGCTCTTCAAAATGCAAGTGCAGTGCAGCTTTTCTGAGCGGCTACTGGGAAATTC 1173
DB 1131 GAGCTGCTCTTCAAAATGCAAGTGCAGTGCAGCTTTTCTGAGCGGCTACTGGGAAATTC 1190

QY 1174 TCATTAGCAGTTCCGATCTTTGTCCTCTCTCTCTGCTTGGCTCCATGAACGGTGGTGTG 1233
DB 1191 TCATTAGCAGTTCCGATCTTTGTCCTCTCTCTCTGCTTGGCTCCATGAACGGTGGTGTG 1250

QY 1234 TTTGCTGCTCCAGGTATTCTATCTGCTGCTGCTGAGAGGTCACCTTCCAGAAATCCTC 1293
DB 1251 TTTGCTGCTCCAGGTATTCTATCTGCTGCTGCTGAGAGGTCACCTTCCAGAAATCCTC 1310

QY 1294 TCATGATTCATCTCGCAAGCACACTCTCTACACAGCTGTTATTGTTTGCACCTTTG 1353
DB 1311 TCATGATTCATCTCGCAAGCACACTCTCTACACAGCTGTTATTGTTTGCACCTTTG 1370

QY 1354 ACAATGATTAATGCTCTCTCTGAGACCTCGACAGTCTTTTGAATTTCTCTGATTTTGGC 1413
DB 1371 ACAATGATTAATGCTCTCTCTGAGACCTCGACAGTCTTTTGAATTTCTCTGATTTTGGC 1430

QY 1414 AGTGGCTTTTATTTGGGCTGGCAGTGTGCTGGCTGATTATCTTCGATACAAATGCCCA 1473
DB 1431 AGTGGCTTTTATTTGGGCTGGCAGTGTGCTGGCTGATTATCTTCGATACAAATGCCCA 1490

QY 1474 GATATCATCTGCTCTTCAAGGTGCCACTGTTTATCCCAAGCTTTGTTTCTTCACATGC 1533
DB 1491 GATATCATCTGCTCTTCAAGGTGCCACTGTTTATCCCAAGCTTTGTTTCTTCACATGC 1550

QY 1534 CTCCTCATGGTGGCTTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1593
DB 1551 CTCCTCATGGTGGCTTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1610

QY 1594 ATCACTCTGACGTGGAGTCCCTGCTGATTAATCTCTTATATATGGGACAAAGAACCCAGG 1653
DB 1611 ATCACTCTGACGTGGAGTCCCTGCTGATTAATCTCTTATATATGGGACAAAGAACCCAGG 1670

QY 1654 TGGTTTGAATTAATGTCAGAGAAATAACAGAAATTAACAATTAATACTGGAAGTTGTA 1713
DB 1671 TGGTTTGAATTAATGTCAGAGAAATAACAGAAATTAACAATTAATACTGGAAGTTGTA 1730

QY 1714 CCAGACAGATTAATGTAATGACATTAATGACATTTGACATCTTGCAATCTGCCCAAGGGA 1773

```

Db 1731 CCAGAGAAGATAAGCTTATGAACCTTAATGACTTGAGATCTTGGAATCTGCCAAGGGGA 1790
QY 1774 GACACAAATAGGGATTTTACTTCATTTCTGAAAGTCTAGAAATACAACTTTGGTG 1833
Db 1791 GACACAAATAGGGATTTTACTTCATTTCTGAAAGTCTAGAAATACAACTTTGGTG 1850
QY 1834 ATAAACAAA 1843
Db 1851 ATAAACAAA 1860

RESULT 7
AB026891 1861 bp mRNA linear PRI 10-FEB-2001
LOCUS Homo sapiens mRNA for cystine/glutamate transporter, complete cds.
DEFINITION AB026891
ACCESSION AB026891
VERSION AB026891.1 GI:5668544
KEYWORDS cystine/glutamate transporter.
SOURCE Homo sapiens fibroblast cdna to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Sato,H., Tamba,M., Kuriyama-Matsumura,K., Okuno,S. and Bannai,S.
Molecular cloning and expression of human xCT, the light chain of
amino acid transporter system xc-
Antioxidants and Redox Signaling 2, 665-671 (2000)
2 (bases 1 to 1861)
Sato,H. and Bannai,S.
Direct Submission
Submitted (30-APR-1999) Hideyo Sato, University of Tsukuba,
Institute of Basic Medical Sciences; Tennodai 1-1-1, Tsukuba,
Ibaraki 305-8575, Japan (E-mail:hideyo-s@md.tsukuba.ac.jp,
Tel:81-298-53-3282, Fax:81-298-53-3039)
FEATURES
source 1..1861
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="fibroblast"
1..1861
/gene="xCT"
236..1741
/gene="xCT"
/codon_start=1
/product="cystine/glutamate transporter"
/protein_id="BAA82628.1"
/db_xref="GI:5668545"
/translation="NVRKPVVSTISKGYLQNVNGLRPLSLNKPPOGKVLKRVV
TLRGVSIILTIIGAGIFIPKGLVNTGVSMSLTITWVGVLSLFCALSYAELGT
TIKSGGHYIILEVFGPLPAFVRVWVLLIIRPAATAVISLAFGRYILEPFIQCEI
PELAIKLITAVGITVVMVLSNVSWSARIQIFLTFCKLAILIIVPGVMOLIKGOT
QNFKAFSGRDSSITRLPLAFYGYMAYAGWFLNFTVEEVENPEKTIPLAICISMAI
VTIGVYLTWAVFTTINABELLSNAVAVTFERLLNFSLAVPLFVALSCFSGNNGG
VFAVSLFVVASREGLHPILLSKIHVRKHTPLPAVIVLPLTMIMFSGDLSLNLFL
SFARWLFGLAVAGLIYLYKCPDMHRFPKVPFLFIPALFSPFCLEFMVALSYSDPFT
GIGFVITLGVPAIYLFILWDRKPRWRIMSEKIRITRLQIILEVVPEDKL"
BASE COUNT 485 a 402 g 421 g 553 t
ORIGIN
Query Match 64.9%; Score 1454; DB 9; Length 1861;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 334 GAGAAAGTCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGCG 393
Db 341 GAGAAAGTCAGCTGAAGAGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGCG 400
QY 394 ACCATCATTCGACAGCAATCTTCATCTCTCTTAAGGGCGTGTCTCAGAACACGGGACG 453
Db 401 ACCATCATTCGACAGCAATCTTCATCTCTCTTAAGGGCGTGTCTCAGAACACGGGACG 460
QY 454 GTGGGATGTCCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCACATTTTGGAGCTTGG 513

```

```

Db 1541 CTCTTCATGTTGCCCTTCTCTATTCCGACCCATTAGTACAGGGATTGGCTCGTC 1600
QY 1594 ATCACTCTGACTGGAGTCCCTGGTATTATCTCTTTATATATGAGCAAGAAACCCAGG 1653
Db 1601 ATCACTCTGACTGGAGTCCCTGGTATTATCTCTTTATATATGAGCAAGAAACCCAGG 1660
QY 1654 TGGTTTGAATAATGTCAGAGAAATTAACAGAAATTAACAATATATCTGGAAGTTGTA 1713
Db 1661 TGGTTTGAATAATGTCAGAGAAATTAACAGAAATTAACAATATATCTGGAAGTTGTA 1720
QY 1714 CCAGAAGAAGATAGTTATGAACATTAATGACTTTGAGACTTTGGCAATCTGCCAATCTGCCAAGGGA 1773
Db 1721 CCAGAAGAAGATAGTTATGAACATTAATGACTTTGAGACTTTGGCAATCTGCCAATCTGCCAAGGGA 1780
QY 1774 GACACAAATAGGGATTTTACTTCTATTTTCGAAAGTCTAGAGATTACAATTTGGTG 1833
Db 1781 GACACAAATAGGGATTTTACTTCTATTTTCTGAAAGTCTAGAGATTACAATTTGGTG 1840
QY 1834 ATAAA 1838
Db 1841 ATAAA 1845

RESULT 8
AB040875
LOCUS
DEFINITION
AB040875 2000 bp mRNA linear PRI 03-APR-2001
Homo sapiens hXCT mRNA for cystine/glutamate exchanger, complete
cds.
AB040875
VERSION
AB040875.1 GI:13516845
Homo sapiens adult cDNA to mRNA, clone_lib:brain cDNA library
clone:hXCT.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Kim,J.Y., Chairoungdua,A., Cha,S.H., Segawa,H., Matsuo,H.,
Kim,D.K., Endou,H. and Kanai,Y.
Human cystine/glutamate exchanger: cDNA cloning and upregulation by
oxidative stress in glioma cells
Unpublished
2 (bases 1 to 2000)
kanai,Y.
Direct Submission
Submitted (01-APR-2000) Yoshikatsu Kanai, Kyorin University School
of Medicine, Department of Pharmacology and Toxicology; 6-20-2
Shinkawa, Mitaka, Tokyo 181-8611, Japan
(E-mail:ykanai@kyorin-u.ac.jp, Tel:+81-422-47-5511(ex.3453),
Fax:+81-422-79-1321)
FEATURES
Location/Qualifiers
1..2000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hXCT"
/clone_lib="brain cDNA library"
/dev_stage="adult"
1..2000
/gene="hXCT"
136..1620
/gene="hXCT"
/feature="Note-independent amino acid transporter
transmembrane protein"
/codon_start=1
/product="cystine/glutamate exchanger"
/protein_id="BAB40574.1"
/db_xref="GI:13516846"
/translocation="MVKRPVVSISKGGYLOGNVNRLPSLGNKEPPQEKVQLRRKY
TLRGVSIITGITIGAGIFISPKGVLONTSGVMSLITWVCGVLSLFGALSTAEIGT
TIKSGGHYITILEVGPPLPAFVRVWVWELLIRPAATAVISLAFGRYILEPFIQCEI
PNAKLITGADGTVVYVNLMSVMSWSARIQIFLTKLTAIILIIIPVGMOLIKQOT
QNKFDAPSRGSSITRLPLAFYGYMAYAGWYLFNFTEVENPENTIPLAICISMAI
VTIGVLTNVAYFTTINAEELLISNAVAVTFSERLILGNFSLAVPIFVALSCFGSMNGG

```

```

VFASRLFYVASREGHLPELLSMHVRKHTPLPAVLVHLPLTMIMLPSGDLSDLLNLF
SFARWFLIGLAVAYLIYLRKCPDMRPFKPLFIIPALFSFCLFVVALSYSDPEST
GIGFVITLTGVPAYLFIWDKKPRFRIMSGFLALMPAQADM"
BASE COUNT 505 a 418 c 435 g 642 t
ORIGIN
Query Match 59.8%; Score 1339; DB 9; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 334 GAGAAAGTGCACCTGAGAGAGAAAGTCACTTACTGAGGGAGTCTCCATATCATTTGGC 393
Db 241 GAGAAAGTGCACCTGAGAGAGAAAGTCACTTACTGAGGGAGTCTCCATATCATTTGGC 300
QY 394 ACCATCATTTGAGAGAGAAATCTTCATCTCTCTAAGGCGCTGCTCCAGAACACGGGCGAG 453
Db 301 ACCATCATTTGAGAGAGAAATCTTCATCTCTCTAAGGCGCTGCTCCAGAACACGGGCGAG 360
QY 454 GTGGGCATGTCCTGACCATCTGGACGGTGTGGGGTCTGTCACATTTTGGAGCTTTG 513
Db 361 GTGGGCATGTCCTGACCATCTGGACGGTGTGGGGTCTGTCACATTTTGGAGCTTTG 420
QY 514 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCAATTACACATATATTTG 573
Db 421 TCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCAATTACACATATATTTG 480
QY 574 GAAAGTCTTTGGTCCATTACCAAGCTTTTGTACGAGTCTGGGTGGAACCTCCATATATACG 633
Db 481 GAAAGTCTTTGGTCCATTACCAAGCTTTTGTACGAGTCTGGGTGGAACCTCCATATATACG 540
QY 634 CTTGAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTGGAACCAATTTT 693
Db 541 CTTGAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTGGAACCAATTTT 600
QY 694 ATTCAATGTGAATCCCTGAACTTCGATCAAGCTCAATACAGCTGTGGGCATTAACGTGA 753
Db 601 ATTCAATGTGAATCCCTGAACTTCGATCAAGCTCAATACAGCTGTGGGCATTAACGTGA 660
QY 754 GTGATGGTCTAAATPAGCATGAGTGTGAGTGGAGCGCCGGATCCAGATTTTCTTAAC 813
Db 661 GTGATGGTCTAAATPAGCATGAGTGTGAGTGGAGCGCCGGATCCAGATTTTCTTAAC 720
QY 814 TTTTGCACAGCTCACAGCAATCTGATAATATAGTCCCTGGAGTTATGACGCTAATATAA 873
Db 721 TTTTGCACAGCTCACAGCAATCTGATAATATAGTCCCTGGAGTTATGACGCTAATATAA 780
QY 874 GGTCAAAACGAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAGGTATTTACGGGTTG 933
Db 781 GGTCAAAACGAGAACTTTAAAGACGCCCTTTTCAGGAAGAGATTCAGGTATTTACGGGTTG 840
QY 934 CCATGGCTTTTATATGGAATGTATGATATGCTGGTGGTGTACCTCAACTTTGTT 993
Db 841 CCATGGCTTTTATATGGAATGTATGATATGCTGGTGGTGTACCTCAACTTTGTT 900
QY 994 ACTGAAGAGTAGAAAACCCCTGAAAAACCAATTCCTCCCTTGAATATGATATCCATGGCC 1053
Db 901 ACTGAAGAGTAGAAAACCCCTGAAAAACCAATTCCTCCCTTGAATATGATATCCATGGCC 960
QY 1054 ATTGTACACCATTTGGCTATGTGTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1113
Db 961 ATTGTACACCATTTGGCTATGTGTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1020
QY 1114 GAGTGTCTCTTCAAAATGACGTGGCAGTACCTTTCTGAGCGGCTACTGGGAAATTTTC 1173
Db 1021 GAGTGTCTCTTCAAAATGACGTGGCAGTACCTTTCTGAGCGGCTACTGGGAAATTTTC 1080
QY 1174 TCATTAGCAGTTCCGATCTTTGTCCTCTCTGCTTTGGCTCCATGAACGGTGGTGTG 1233
Db 1081 TCATTAGCAGTTCCGATCTTTGTCCTCTCTGCTTTGGCTCCATGAACGGTGGTGTG 1140
QY 1234 TTTGCTGTCTCCAGGTATTCATGTTGGCTGTGAGAGGGTTCACCTTCCAGAAATCTTC 1293
Db 1141 TTTGCTGTCTCCAGGTATTCATGTTGGCTGTGAGAGGGTTCACCTTCCAGAAATCTTC 1200

```

QY 1294 TCCATGATTCATGTCGCAAGCACACTCCTCTACACAGCTGTATATGTTTTCACCCCTTTG 1353
 |||||
 Db 1201 TCCATGATTCATGTCGCAAGCACACTCCTCTACACAGCTGTATATGTTTTCACCCCTTTG 1260
 |||||
 QY 1354 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTGAATTCCTCAGTTTGGCC 1413
 |||||
 Db 1261 ACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTGAATTCCTCAGTTTGGCC 1320
 |||||
 QY 1414 AGGTGGCTTTTATTTGGGCTGGCAGTCTGCTGGCTGATTTATCTTCGATACAAATGCCCA 1473
 |||||
 Db 1321 AGGTGGCTTTTATTTGGGCTGGCAGTCTGCTGGCTGATTTATCTTCGATACAAATGCCCA 1380
 |||||
 QY 1474 GATATGCATCGCTCTTCAAGGTGCCACTGTTCATCCAGCTTTTGTTCCTTCACATGCG 1533
 |||||
 Db 1381 GATATGCATCGCTCTTCAAGGTGCCACTGTTCATCCAGCTTTTGTTCCTTCACATGCG 1440
 |||||
 QY 1534 CTTCTCATGTTGGCCCTTCCCTCTATTCGACCCATTTAGTACAGGATTTGGCTTCGTC 1593
 |||||
 Db 1441 CTTCTCATGTTGGCCCTTCCCTCTATTCGACCCATTTAGTACAGGATTTGGCTTCGTC 1500
 |||||
 QY 1594 ATCACTCTGACTGGAGTCCCTCGCTATTTATCTCTTTATATGCGACAAACCCAGG 1653
 |||||
 Db 1501 ATCACTCTGACTGGAGTCCCTCGCTATTTATCTCTTTATATGCGACAAACCCAGG 1560
 |||||
 QY 1654 TGGTTTAGAATAATGTCAG 1672
 |||||
 Db 1561 TGGTTTAGAATAATGTCAG 1579
 |||||

RESULT 9
 AC110804/c
 LOCUS Homo sapiens BAC clone RP11-725C19 from 4, complete sequence.
 DEFINITION AC110804
 ACCESSION AC110804
 VERSION AC110804.3 GI:19570222
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 64196)
 Sulston, J.E. and Waterston, R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 99063792
 PUBMED 9847074
 2 (bases 1 to 64196)
 Pearman, C., Kozlowicz, A. and Dignan, G.
 The sequence of Homo sapiens BAC clone RP11-725C19
 Unpublished (2001)
 3 (bases 1 to 64196)
 Waterston, R.H.
 Direct Submission
 Submitted (15-FEB-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 64196)
 Waterston, R.H.
 Direct Submission
 Submitted (06-MAR-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 64196)
 Waterston, R.H.
 Direct Submission
 Submitted (21-MAR-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 6 (bases 1 to 64196)
 Waterston, R.
 Direct Submission
 Submitted (30-MAR-2002) Department of Genetics, Washington

COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 21, 2002 this sequence version replaced gi:19073866.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.edu
 ----- Summary Statistics

 Center project name: H_NH0725C19

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org
 VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-8p17, 2000 bp overlap.
 Actual start of this clone is at base position 185268 of RP11-393A12; actual end is at base position 64196 of RP11-725C19.

Polymorphisms exist between AC013365 and AC110804.

FEATURES
 source
 Location/Qualifiers
 1..64196
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /map="4"
 /clone="RP11-725C19"
 /clone_lib="RPC1-11"
 1..39
 /rpt_family="AT_rich"
 162..182
 /rpt_family="AT_rich"
 312..607
 /rpt_family="Alu"
 369..505
 /note="match to EST BG257959 (NID:gl2767775)"
 877..909
 /rpt_family="AT_rich"
 1113..1536
 /rpt_family="L2"
 2400..2427
 /rpt_family="AT_rich"
 2941..2984
 /rpt_family="L2"
 3413..3435
 /rpt_family="(TA)n"
 3829..4216
 misc_feature

misc_feature /note="match to EST H06135 (NID:g869687) y177b12.s1"
3829.4170
/note="match to EST B1497113 (NID:g15336457)"
misc_feature 3831.4157
/note="match to EST A639698 (NID:g2563477) np60f09.s1"
misc_feature 3835.4360
/note="match to EST AW237145 (NID:g6569534) xm66b03.x1"
misc_feature 3836.4376
/note="similar to EST BM791320 (NID:g19139552)"
misc_feature 3837.4405
/note="similar to Bos taurus EST B1539797 (NID:g15380907)"
misc_feature 3844.4610
/note="match to EST B1857800 (NID:g15998547)"
misc_feature 3933.4394
/note="match to EST A1694383 (NID:g4971723) wd83b04.x1"
misc_feature 4105.4843
/note="match to EST BGL177619 (NID:g12684322)"
misc_feature 4137.4527
/note="match to EST A1267688 (NID:g3886855) aq93c10.x1"
misc_feature 4153.4433
/note="match to EST C00538 (NID:g1432768)"
misc_feature 4233
/note="match to EST B1497113 (NID:g15336457)"
misc_feature 4359.4746
/note="match to EST AA227760 (NID:g1849333) zr56a07.r1"
misc_feature 4404.5178
/note="match to EST BG291471 (NID:g13049386)"
misc_feature 4412.4619
/note="similar to Homo sapiens EST BF154977 (NID:g11050160)"
misc_feature 4657.5028
/note="match to EST R20307 (NID:g774941) yg35a09.r1"
misc_feature 4727.5195
/note="similar to Bos taurus EST AW669076 (NID:g7525590)"
misc_feature 4740.5803
/note="similar to EST BM547723 (NID:g18781736)"
misc_feature 4859.5265
/note="match to EST AW901105 (NID:g8065310)"
misc_feature 5002.5687
/note="match to EST BG254117 (NID:g12763933)"
misc_feature 5004.5213
/note="match to EST BF084219 (NID:g10878049)"
misc_feature 5074.5447
/note="match to EST N80420 (NID:g1243121) yz92e06.r1"
misc_feature 5255.6002
/note="match to EST B1869486 (NID:g16043159)"
misc_feature 5442.6151
/note="match to EST AW961203 (NID:g8150887)"
misc_feature 5497.5830
/note="match to EST A1540897 (NID:g4458270)"
misc_feature 5782.6124
/note="match to EST AA309186 (NID:g1961511)"
misc_feature 5875.6701
/note="match to EST B1334972 (NID:g15019629)"
misc_feature 5995.6331
/note="match to EST AA471165 (NID:g2199022)"
misc_feature 6000.6124
/note="match to EST AW996896 (NID:g8257130)"
misc_feature 6032.6836
/note="similar to EST BM474943 (NID:g18523985)"
misc_feature 6140.6518
/note="match to EST AW996897 (NID:g8257131)"
misc_feature 6444.6885
/note="match to EST AA375957 (NID:g2028275)"
repeat_region 6878.7187
/rpt_family="Alu"
misc_feature 6899.7362
/note="match to EST AA488687 (NID:g2216118) ab38f03.s1"
misc_feature 6970.7746
/note="match to EST BE739933 (NID:g10153837)"
misc_feature 7099.7815
/note="match to EST BF381953 (NID:g11363256)"
misc_feature 7124.7125

misc_feature /note="match to EST AW996897 (NID:g8257131)"
7174.7495
/note="match to EST BF927528 (NID:g12325656)"
misc_feature 7253.7884
/note="similar to Homo sapiens EST BG499179 (NID:g13460696)"
misc_feature 7568.7949
/note="match to EST AA307001 (NID:g1959330)"
misc_feature 7613.8332
/note="similar to EST BM706727 (NID:g19019985)"
repeat_region 7836.8011
/rpt_family="AcHobo"
misc_feature 8035.8572

Query Match 21.4%; Score 479; DB 9; Length 64196;
Best Local Similarity 99.8%; Pred. No. 4.5e-261;
Matches 529; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1671 AGAGAAATACAGAACATTACAATAATCTGGAAGTTGTACAGAGAGAGATAAGTT 1730
|||||
Db 11755 AGAGAAATACAGAACATTACAATAATCTGGAAGTTGTACAGAGAGAGATAAGTT 11696
|||||

QY 1731 ATGAATATGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAATAGGGATT 1790
|||||
Db 11695 ATGAATATGACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAATAGGGATT 11636
|||||

QY 1791 TTTACTTCAATTTCTGAAAGCTAGAGAAATTACAATTTGGTGATAACAAAGAGGTCA 1850
|||||
Db 11635 TTTACTTCAATTTCTGAAAGCTAGAGAAATTACAATTTGGTGATAACAAAGAGGTCA 11576
|||||

QY 1851 GTTATTTTATTCATATATTTTAGCATATTCGAACATAATTTCTAAGAAATTTAGTTATAA 1910
|||||
Db 11575 GTTATTTTATTCATATATTTTAGCATATTCGAACATAATTTCTAAGAAATTTAGTTATAA 11516
|||||

QY 1911 CTCTATGATTTATGAAAGTGAATATGCAATTTCTTATGATCGCACAAATCTTGAGT 1970
|||||
Db 11515 CTCTATGATTTATGAAAGTGAATATGCAATTTCTTATGATCGCACAAATCTTGAGT 11456
|||||

QY 1971 CTCTATACCTACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTGGTCATTC 2030
|||||
Db 11455 CTCTATACCTACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTGGTCATTC 11396
|||||

QY 2031 TCTCAACATATCTTAGCAGCGCAAGAACCTTCAAAATGAAGACTGAGATTTTCTGTA 2090
|||||
Db 11395 TCTCAACATATCTTAGCAGCGCAAGAACCTTCAAAATGAAGACTGAGATTTTCTGTA 11336
|||||

QY 2091 TATATGGGTTTGTAAAGATGTTTTTACACACTACAGATGTCTATCTGTGAAAAGTGT 2150
|||||
Db 11335 TATATGGGTTTGTAAAGATGTTTTTACACACTATAGATGTCTATCTGTGAAAAGTGT 11276
|||||

QY 2151 TTCAATCTGAAAAAAGCATACATCATGATTATGCAAGAGGAGAGAA 2200
|||||
Db 11275 TTCAATCTGAAAAAAGCATACATCATGATTATGCAAGAGGAGAGAA 11226
|||||

RESULT 10
AX321911
LOCUS AX321911
DEFINITION Sequence 442 from Patent WO0172295.
ACCESSION AX321911
VERSION AX321911.1 GI:17906521
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
Mannion,J. and Kalos,M.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0172295-A 442 04-OCT-2001;

FEATURES
source CORIXA CORPORATION (US)
Location/Qualifiers
1..337
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 103 a 60 c 93 g 81 t
ORIGIN

Query Match 12.6%; Score 281; DB 6; Length 337;
Best Local Similarity 100.0%; Pred. No. 4.1e-148;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAGGTGAAGTGACGAGAGATCATGCCAGCCTGGGTGACAGTGTCTGTCTCAAAAC 60
Db 4 GGAGGTGAAGTGACGAGAGATCATGCCAGCCTGGGTGACAGTGTCTGTCTCAAAAC 63
QY 61 ACAATTAAGGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAAATCCAGGCCAATTGTGG 120
Db 64 AGAATTAAGGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAAATCCAGGCCAATTGTGG 123
QY 121 CATAGATTTTATCATATTTCTGGATTTTGGATTTCTTTTGTCTCATCTACTGGATTCA 180
Db 124 CATAGATTTTATCATATTTCTGGATTTTGGATTTCTTTTGTCTCATCTACTGGATTCA 183
QY 181 GGAAGCCTGTGTCTCCACCATCTCCAAAGGAGTTACCTGCAGGGAAATGTTAACGGG 240
Db 184 GGAAGCCTGTGTCTCCACCATCTCCAAAGGAGTTACCTGCAGGGAAATGTTAACGGG 243
QY 241 AGGCTGCCTTCCCTGGGCAACAGGAGCCACTGGGCAGGA 281
Db 244 AGGCTGCCTTCCCTGGGCAACAGGAGCCACTGGGCAGGA 284

RESULT 11
AX351341/c
LOCUS AX351341 596 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 88 from Patent WO0196390.
ACCESSION AX351341
VERSION AX351341.1 GI:18616688
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
Jiang, Y., Hepler, W.T., Clapper, J.D., Wang, A. and Secrist, H.
Compositions and methods for the therapy and diagnosis of colon
cancer
Patent: WO 0196390-A 88 20-DEC-2001;
CORIXA CORPORATION (US)
Location/Qualifiers
1..596
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 160 a 151 c 126 g 158 t 1 others
ORIGIN

Query Match 12.0%; Score 269; DB 6; Length 596;
Best Local Similarity 100.0%; Pred. No. 2.9e-141;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAAGTGCAGCTGAAGAGGAAAGTCTACTTACTAGGGGAGTCTCATTATCATTTGGC 393
Db 269 GAGAAGTGCAGCTGAAGAGGAAAGTCTACTTACTAGGGGAGTCTCATTATCATTTGGC 210
QY 394 ACCATCATTTGGAGCAGGAATCTTCTCTCTTAAGGCGTCTCCAGAACACGGGAGC 453
Db 209 ACCATCATTTGGAGCAGGAATCTTCTCTCTTAAGGCGTCTCCAGAACACGGGAGC 150
QY 454 GTGGCGATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTCACTATTTGGAGCTTG 513
Db 149 GTGGCGATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTCACTATTTGGAGCTTG 90

QY 514 TCTTATGCTGAATTTGGGAACAACATATATAAGAAATCTGGAGTCAATACACATATATTTG 573
Db 89 TCTTATGCTGAATTTGGGAACAACATATATAAGAAATCTGGAGTCAATACACATATATTTG 30
QY 574 GAAGTCTTTGGTCCATTACCAGCTTTTGT 602
Db 29 GAAGTCTTTGGTCCATTACCAGCTTTTGT 1

RESULT 12
AC093903/c
LOCUS AC093903 161280 bp DNA linear PRI 01-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-733C7 from 4, complete sequence.
ACCESSION AC093903 AC055827
VERSION AC093903.3 GI:15920156
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 161280)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
REFERENCE 2 (bases 1 to 161280)
AUTHORS Radonenko, M. and Kozlowicz, A.
TITLE The sequence of Homo sapiens BAC clone RP11-733C7
JOURNAL Unpublished (2001)
MEDLINE 99063792
REFERENCE 3 (bases 1 to 161280)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 161280)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 161280)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 4, 2001 this sequence version replaced gi:15625016.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0733C07
Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.

McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/qsc>

SOURCE INFORMATION:

SOURCE INFORMATION:
The RPlci-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanesi, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletzer de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is RP11-310A13. Actual start of this clone is at base position 1 of RP11-733C7; actual end is at base position 161280 of RP11-733C7.

Data from AC009792 was used to finish the clone. AC055827.

The sequence of AC055827 has been incorporated into AC093903.

```

Location/Qualifiers
1. .161280
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"

```

repeat_region	/clone_lib="APC1-L1"	148. .304
repeat_region	/rpt_family="L1"	306. .406
repeat_region	/rpt_family="(GGAA)n"	471. .525
repeat_region	/rpt_family="AT-rich"	530. .782
repeat_region	/rpt_family="L1"	685. .712
repeat_region	/rpt_family="(TA)n"	1383. .1472
repeat_region	/rpt_family="CA-rich"	1697. .1836
repeat_region	/rpt_family="L1"	1837. .1956
repeat_region	/rpt_family="L1"	3005. .3221
repeat_region	/rpt_family="MIR"	4013. .4321
repeat_region	/rpt_family="Alu"	4295. .4325
repeat_region	/rpt_family="(A)n"	5625. .5960
repeat_region	/rpt_family="L1"	5938. .5982
repeat_region	/rpt_family="A-rich"	6062. .6195
repeat_region	/rpt_family="MIR"	6388. .6462
repeat_region	/rpt_family="L2"	6546. .6868
misc_feature	/rpt_family="L1"	6866. .7318
repeat_region	/note="similar to EST"	6866. .6898
misc_feature	/rpt_family="AT-rich"	7134. .7395
misc_feature	/note="similar to EST"	7777. .7918
repeat_region	/note="similar to EST"	8483. .8522
repeat_region	/rpt_family="AT-rich"	8721. .9018

/rpt_family
8721. .9018

repeat_region	/rpt_family="Alu" 9459. .9741	/note="similar to EST BG388107 (NID:g13281553)"
repeat_region	/rpt_family="I2" 10217. .10246	/note="similar to EST BF286052 (NID:g11217122)"
repeat_region	/rpt_family="AT_rich" 10304. .10385	/note="similar to EST BF095483 (NID:g10901193)"
repeat_region	/rpt_family="MIR" 10835. .10878	/note="similar to EST BG181259 (NID:g13703031)"
misc_feature	/rpt_family="AT_rich" 11861. .11987	/note="similar to EST BG388107 (NID:g13281553)"
misc_feature	/note="similar to EST BF286052 (NID:g11217122)" 11873. .11987	/note="similar to EST BF095483 (NID:g10901193)"
repeat_region	/rpt_family="TA)n" 12516. .12539	/note="similar to EST BG181259 (NID:g13703031)"
repeat_region	/rpt_family="Alu" 12929. .13232	/note="similar to EST BF286052 (NID:g11217122)"
repeat_region	/rpt_family="MIR" 13237. .13274	/note="similar to EST BG388107 (NID:g13281553)"
repeat_region	/rpt_family="TA)n" 13275. .13304	/note="similar to EST BF095483 (NID:g10901193)"
repeat_region	/rpt_family="Alu" 13696. .13838	/note="similar to EST BG181259 (NID:g13703031)"
repeat_region	/rpt_family="MIR" 14127. .14153	/note="similar to EST BF286052 (NID:g11217122)"
repeat_region	/rpt_family="AT_rich" 14132. .14436	/note="similar to EST BG388107 (NID:g13281553)"
repeat_region	/rpt_family="Alu" 15823. .16117	/note="similar to EST BF095483 (NID:g10901193)"
repeat_region	/rpt_family="Alu" 17329. .17894	/note="similar to EST BG181259 (NID:g13703031)"
misc_feature	/note="similar to EST BG388107 (NID:g13281553)" 17329. .17618	/note="similar to EST BF286052 (NID:g11217122)"
misc_feature	/note="similar to EST BF286052 (NID:g11217122)" 17489. .17611	/note="similar to EST BG388107 (NID:g13281553)"
misc_feature	/note="similar to EST AI313891 (NID:g4029010)" 17548. .17869	/note="similar to EST BF095483 (NID:g10901193)"
misc_feature	/note="similar to EST BF095483 (NID:g10901193)" 17772. .18329	/note="similar to EST BG181259 (NID:g13703031)"
repeat_region	/note="similar to EST BG181259 (NID:g13703031)" 17845. .17869	/note="similar to EST BF286052 (NID:g11217122)"
repeat_region	/rpt_family="CTG)n" 18792. .18828	/note="similar to EST BG388107 (NID:g13281553)"
repeat_region	/rpt_family="AT_rich" 20734. .20784	/note="similar to EST BF095483 (NID:g10901193)"
repeat_region	/rpt_family="MERL_type" 20832. .20944	/note="similar to EST BG181259 (NID:g13703031)"
repeat_region	/rpt_family="Alu" 20948. .21964	/note="similar to EST BF286052 (NID:g11217122)"
repeat_region	/rpt_family="L1" 21964. .21986	/note="similar to EST BG388107 (NID:g13281553)"
repeat_region	/rpt_family="AT_rich" 21965. .22276	/note="similar to EST BF095483 (NID:g10901193)"
repeat_region	/rpt_family="Alu" 22277. .22323	/note="similar to EST BG181259 (NID:g13703031)"
repeat_region	/rpt_family="L1" 22341. .22681	/note="similar to EST BF286052 (NID:g11217122)"
repeat_region	/rpt_family="L1" 22685. .22824	/note="similar to EST BG388107 (NID:g13281553)"
repeat_region	/rpt_family="Alu" 22805. .22824	/note="similar to EST BF095483 (NID:g10901193)"
repeat_region	/rpt_family="A)n" 22826. .23575	/note="similar to EST BG181259 (NID:g13703031)"
repeat_region	/rpt_family="L1" 23591. .23735	/note="similar to EST BF286052 (NID:g11217122)"
repeat_region	/rpt_family="MERL_type" 23591. .23735	/note="similar to EST BG388107 (NID:g13281553)"

Query Match 8.1%; Score 182; DB 9; Length 161280;
Best Local Similarity 100.0%; Pred. No. 1.6e-91;
Matches 182; Conservative 0; Mismatches 0; Indels 0;

Qy 1 GGAGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAC 60
|||||
Db 27756 GGAGTTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAC 27697

QY 61 AGAATTAAGGAAAGAAAGAAAGAAAGAGAGAGAGAAATTCAGGCCAATTGTGG 120
 |||||
 Db 27696 AGAATTAAGGAAAGAAAGAAAGAAAGAGAGAGAGAAATTCAGGCCAATTGTGG 27637
 |||||
 QY 121 CATAGATTTATCATATTTCTGGATTTTGTGGATTTCTTTTCTCATCACTGGATTC 180
 |||||
 Db 27636 CATAGATTTATCATATTTCTGGATTTTGTGGATTTCTTTTCTCATCACTGGATTC 27577
 |||||
 QY 181 GG 182
 ||
 Db 27576 GG 27575

RESULT 13
 LOCUS AB042201 1683 bp DNA linear PRI 10-FEB-2001
 DEFINITION Homo sapiens xCT gene for cystine/glutamate transporter, exon 1, partial cds.
 ACCESSION AB042201
 VERSION AB042201.1 GI:7670267
 KEYWORDS cystine/glutamate transporter.
 SOURCE Homo sapiens DNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)
 Sato.H., Tamba.M., Kuriyama-Matsumura.K., Okuno.S. and Bannai.S.
 TITLE Molecular cloning and expression of human xCT, the light chain of amino acid transport system xc-
 JOURNAL Antioxidants and Redox Signaling 2, 665-671 (2000)
 REFERENCE 2 (bases 1 to 1683)
 AUTHORS Sato.H. and Bannai.S.
 TITLE Direct Submission
 JOURNAL Submitted (24-APR-2000) Hideyo Sato, University of Tsukuba, Institute of Basic Medical Sciences; Tennodai 1-1-1, Tsukuba, Ibaraki 305-8575, Japan (E-mail:hideyo-send.tsukuba.ac.jp, Tel:81-298-53-3282, Fax:81-298-53-3039)
 FEATURES
 source 1..1683
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1001..1683
 /gene="xCT"
 1001..1555
 /gene="xCT"
 /product="cystine/glutamate transporter"
 /number=1
 1279..>1555
 /gene="xCT"
 /codon_start=1
 /product="cystine/glutamate transporter"
 /protein_id="BAA94999.1"
 /db_xref="GI:7670268"
 translation="MVRKPVSTISKGYLQGNVNRPLSLNKKPPGQEKVQLRKRV
 TLLRGVSIIGITIGAGIFISPKVLQNTGSVGMSLTIWTCGVLSLF"
 1556..>1683
 /gene="xCT"
 /number=1

BASE COUNT 495 a 287 c 370 g 531 t
 ORIGIN

Query Match 7.7%; Score 173; DB 9; Length 1683;
 Best Local Similarity 100.0%; Pred. No. 1.9e-86;
 Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGAACTGACGTGAAGAGGAAGTCACTTTACTGAGGGAGTCCTCCATTATCATTTGGC 393
 |||||
 Db 1384 GAGAAAGTGCAGCTGAAGAGGAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 1443
 |||||
 QY 394 ACCATCATTTGACGAGGAATTCATCTCCTTAAGGGCGTCTCCAGACACAGGGCAGC 453
 |||||
 Db 1444 ACCATCATTTGACGAGGAATTCATCTCCTTAAGGGCGTCTCCAGACACAGGGCAGC 1503
 |||||

QY 454 GTGGCATGTCTGTGACCATCTGGACGGTGTGTGGGGTCTGTCTGTCATATTGG 506
 |||||
 Db 1504 GTGGCATGTCTGTGACCATCTGGACGGTGTGTGGGGTCTGTCTGTCATATTGG 1556
 |||||

RESULT 14
 LOCUS AC116610 3994 bp DNA linear PRI 29-MAY-2002
 DEFINITION Homo sapiens BAC clone RP13-562H2 from 4, complete sequence.
 ACCESSION AC116610
 VERSION AC116610.4 GI:20279502
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3994)
 Sulston,J.E. and Waterston,R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 3994)
 AUTHORS Desai.A. and Cotton.M.
 TITLE The sequence of Homo sapiens BAC clone RP13-562H2
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 3994)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 3994)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (24-APR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 3994)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAY-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Apr. 24, 2002 this sequence version replaced gi:20136935.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.edu
 ----- Summary Statistics

 Center project name: H_FH0562H02

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCT-13 Human Female BAC Library was constructed using improved cloning techniques developed by Kazutoyo Osoegawa. The library was generated by Baohui Zhao in our laboratory. Construction was funded by a grant from the National Human Genome Research Institute (NHGRI, NIH) (#1R01HG01165-03). The library was generated according to the new NHGRI/DOE 'Guidance on Human Subjects in Large-Scale DNA Sequencing'.

Female blood was obtained via a double-blind selection protocol. Female blood DNA was isolated from one randomly chosen donor (out of 10 female donors) and partially digested with a combination of EcoRI and EcoRI Methylase for library segments 1&2 or either MboI or DpnII for library segments 3&4. Size selected DNA was cloned into the pBAC3.6 vector between the EcoRI sites for library segments 1&2 or the BamHI sites for library segments 3&4. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). The library has been arrayed into 384-well microtiter dishes and also gridded onto 22x22cm nylon high density filters for screening by probe hybridization.

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-725C19, 2000 bp overlap; the clone sequenced to the right is RP11-733C7, 2000 bp overlap. Actual start of this clone is at base position 92698 of RP11-8P17; actual end is at base position 68420 of RP11-733C7.

A single plasmid subclone region exists between bases 2382-2397.

Polymorphisms have been identified between AC110804, AC093903, AC105310, and AC116610.

FEATURES

```

source
1..3994
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP13-562H2"
/clone_lib="RPCT-13"
169..205
/rpt_family="AT-rich"
935..1089
/rpt_family="MER1_type"
1606..1716
/rpt_family="L2"
2142..2298
/rpt_family="L1"
2300..2400
/rpt_family="(GGAA)n"
2465..2519
/rpt_family="AT-rich"
2524..2678
/rpt_family="L1"
2679..2716
/rpt_family="(TA)n"
2717..2786
/rpt_family="L1"
3387..3476
/rpt_family="GA-rich"
3701..3840
/rpt_family="L1"
3841..3960
/rpt_family="L1"
BASE COUNT 1431 a 656 c 689 g 1218 t
ORIGIN

```

Query Match 5.7%; Score 128; DB 9; Length 3994;
 Best Local Similarity 100.0%; Pred. No. 9.5e-61;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 749 CTGTAGTGTGGTCTCTAAATAGCATGATGTCAGTCGAGCGCCCGGATCCAGATTTCCT 808
|||||
DB 864 CTGTAGTGTGGTCTCTAAATAGCATGATGTCAGTCGAGCGCCCGGATCCAGATTTCCT 805
|||||
QY 809 TAACTTTTGGCAAGCTCACACCAATTCTGATAATTATATAGTCCCTGGAGTTATGCAGCTAA 868
|||||
DB 804 TAACTTTTGGCAAGCTCACACCAATTCTGATAATTATATAGTCCCTGGAGTTATGCAGCTAA 745
|||||
QY 869 TTAAGGT 876
|||||
DB 744 TTAAGGT 737
|||||

```

RESULT 15

```

AX341060 LOCUS AX341060 575 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 1307 from Patent WO0196388.
ACCESSION AX341060
VERSION AX341060.1 GI:18137042
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE

```

1
AUTHORS Jiang,Y., Harlocker,S.L. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon
cancer
JOURNAL Patent: WO 0196388-A 1307 20-DEC-2001;
CORIXA CORPORATION (US)

```

FEATURES

source

```

1..575
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 151 a 122 c 153 g 141 t 8 others
ORIGIN

```

Query Match 4.8%; Score 107; DB 6; Length 575;
 Best Local Similarity 100.0%; Pred. No. 8.8e-49;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 334 GAGAAAGTGCAGCTCAAGAGGAAGTCTTACTGAGGGAGTCTCATTTATCATTTATCATTTGGC 393
|||||
DB 354 GAGAAAGTGCAGCTCAAGAGGAAGTCTTACTGAGGGAGTCTCATTTATCATTTATCATTTGGC 413
|||||
QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGCGGTGCTCCA 440
|||||
DB 414 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGCGGTGCTCCA 460
|||||

```

Search completed: April 16, 2003, 13:16:27

Job time : 6862 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 10:51:38 ; Search time 3547 Seconds
(without alignments)
10223.201 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgaagtgcagag.....ttattaaaaaaaaaaaaa 2239

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpi: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_othe: *
26: em_gss_pro: *
27: em_gss_rod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	455	20.3	516	9	AL702210
c	450	20.1	483	10	AW205572
c	438	19.6	511	14	BM887840
4	429	19.2	919	12	BG490131
5	262	11.7	910	12	BG326527
6	236	10.5	936	12	BG284503

c	7	172	7.7	557	17	AQ002318
	8	171	7.6	371	14	W00842
	9	153	6.8	400	17	AQ533681
	10	153	6.8	486	17	B33331
	11	152	6.8	497	14	BQ370588
	12	128	5.7	485	17	BH860963
	13	112	5.0	1072	12	BG388107
	14	102	4.6	752	12	BG186884
c	15	87	3.9	87	10	AW059922
c	16	84	3.8	632	17	AQ201392
	17	78	3.5	388	9	AA812721
c	18	52	2.3	530	9	AI102370
c	19	51	2.3	123	9	AI133891
c	20	50	2.2	869	10	BE535700
c	21	49	2.2	235	14	BQ448503
c	22	49	2.2	255	10	AW449544
c	23	49	2.2	417	13	BM662461
c	24	49	2.2	602	13	BM511762
c	25	49	2.2	627	10	AV704488
c	26	49	2.2	668	14	BM973417
	27	49	2.2	688	10	AV706925
	28	49	2.2	1028	14	BQ428200
	29	45	2.0	179	12	BF031033
c	30	45	2.0	237	12	BE857736
c	31	45	2.0	262	14	N72154
c	32	45	2.0	270	14	F15607
c	33	45	2.0	286	9	AI125945
c	34	45	2.0	286	9	AI375209
c	35	45	2.0	324	9	AA722121
c	36	45	2.0	343	9	AA922752
c	37	45	2.0	343	9	AI123247
c	38	45	2.0	350	12	BF095483
c	39	45	2.0	358	9	AA398455
c	40	45	2.0	359	14	F37452
c	41	45	2.0	364	9	AI424212
c	42	45	2.0	372	14	BQ013223
c	43	45	2.0	402	14	BQ017470
c	44	45	2.0	409	9	AA922074
c	45	45	2.0	422	9	AI753931

ALIGNMENTS

RESULT 1
AL702210 516 bp mRNA linear EST 22-MAR-2002
LOCUS DKFZp686C09156_r1 686 (synonym: hicc3) Homo sapiens cDNA clone
DEFINITION DKFZp686C09156 5', mRNA sequence.

ACCESSION AL702210 GI:19685565

VERSION EST.

KEYWORDS human.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 516)
Boecker, H., Boecker, M., Brandt, P., Mewes, W., Weil, B. and Wiemann, S.

EST (Boecker, H., Boecker, M., Brandt, P., Mewes, W., Weil, B. and Wiemann, S.)
Unpublished (1999)
Contact: Boecker H

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by GBF (National Research Centre for Biotechnology Ltd.,

Braunschweig/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No sl sequence available.

This clone (DKFZp686C09156) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1. .516
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686C09156"
/tissue="adult"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiI; B;
cDNA-collection"

BASE COUNT 107 a 121 c 110 g 178 t

ORIGIN

Query Match 20.3%; Score 455; DB 9; Length 516;
Best Local Similarity 100.0%; Pred. No. 6.1e-125; Indels 0; Gaps 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 AAAACCTGAAAAACCATTCCTCCATATGATATCCATGCGCATTTGCACCATG 1066
|||||
DB 62 AAAACCTGAAAAACCATTCCTCCATATGATATCCATGCGCATTTGCACCATG 121
|||||

QY 1067 GCTATGCTGCAAAATGCGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTCTT 1126
|||||
DB 122 GCTATGCTGCAAAATGCGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTCTT 181
|||||

QY 1127 CAATGAGTGGCAGTACCTTTCTGAGCGGTACTGGAAATTCATAGCAGTTC 1186
|||||
DB 182 CAATGAGTGGCAGTACCTTTCTGAGCGGTACTGGAAATTCATAGCAGTTC 241
|||||

QY 1187 CGATCTTTGTCCTCTCTGCTTGGCTCCATGACGCTGGTGTGTTGCTGCTCCA 1246
|||||
DB 242 CGATCTTTGTCCTCTCTGCTTGGCTCCATGACGCTGGTGTGTTGCTGCTCCA 301
|||||

QY 1247 GGTATTTCTATGTTGGCTCTCGAGAGGTCACCTCCAGAAATCCCTCCATGATTCATG 1306
|||||
DB 302 GGTATTTCTATGTTGGCTCTCGAGAGGTCACCTCCAGAAATCCCTCCATGATTCATG 361
|||||

QY 1307 TCCGACGACACCTCTCTACGAGCTGTATTTGTCACCTTTGACAAATGATATGC 1366
|||||
DB 362 TCCGACGACACCTCTCTACGAGCTGTATTTGTCACCTTTGACAAATGATATGC 421
|||||

QY 1367 TCTTCTGAGACCTCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGCTTTTGA 1426
|||||
DB 422 TCTTCTGAGACCTCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGCTTTTGA 481
|||||

QY 1427 TTGGCTGGCAGTTGCTGGGCTGATTTATCTTCA 1461
|||||
DB 482 TTGGCTGGCAGTTGCTGGGCTGATTTATCTTCA 516
|||||

RESULT 2

AW205572/c 483 bp mRNA linear EST 02-DEC-1999
LOCUS
DEFINITION
IMAGE:2718171 3', mRNA sequence.

ACCESSION
AW205572
VERSION
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 483)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicg.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=yes.

FEATURES

source

Location/Qualifiers
1. .483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2718171"
/clone_lib="NCI-CGAP_Sub3"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pTriplex2; Site_1: Not I; Site_2: Eco RI; The
NCI-CGAP_Sub3 library is a subtraced library derived from
the NCI-CGAP_Sub1 library, which is a subtraced library
derived from B1. B1 constitutes a mixture of 21
normalized or subtraced NCI-CGAP libraries: NCI-CGAP_Co10,
NCI-CGAP_Pr22, NCI-CGAP_Kid5, NCI-CGAP_Kid12,
NCI-CGAP_Kid3, NCI-CGAP_Kid11, NCI-CGAP_Lym2,
NCI-CGAP_Br2, NCI-CGAP_Co8, NCI-CGAP_CLL1, NCI-CGAP_Le12,
NCI-CGAP_Brn23, NCI-CGAP_Lu5, NCI-CGAP_Lu24,
NCI-CGAP_Lu19, NCI-CGAP_GC4, NCI-CGAP_GC6,
NCI-CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:
NCI-CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE Clonoids 1322376-1323911, 1456008-1456775
1500552-1502855); NCI-CGAP_Kid5 pool 1 LLAM 3338-3342
3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831,
1471368-1472903, 1492104-1493255); NCI-CGAP_Lu5 pool 1
LLAM 3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991,
1520904-1522439); NCI-CGAP_GC4 pool 1 LLAM 3164-3167,
3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631,
1469064-1470983, 1475592-1476743); NCI-CGAP_Pr22 pool 1
LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids
985608-986759, 1101192-1101959, 1217928-1220615);
NCI-CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
Clonoids 1057416-1061255, 1144584-1145351). Subtraction
was performed as previously described [Bonald, Lennon &
Soares (1996): Normalization and Subtraction: Two
Approaches To Facilitate Gene Discovery. Genome Research
6, 791-806.

TAG_LIB=NCI-CGAP_Co10

TAG_TISSUE=colon

TAG_SEQ=AAACG

BASE COUNT 150 a 98 c 98 g 137 t

ORIGIN

Query Match 20.18; Score 450; DB 10; Length 483;
Best Local Similarity 100.0%; Pred. No. 1.9e-123;
Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1394 TGAATTTCTCAGTTTCCAGGTGGCTTTTATTGGCTGGCAGTTCCTGGGCTGATT 1453
|||||
DB 459 TGAATTTCTCAGTTTCCAGGTGGCTTTTATTGGCTGGCAGTTCCTGGGCTGATT 400
|||||

QY 1454 ATCTTCGATACAAATGCCAGATATGCATCGTCCTTTCAAGTCCCACTGTTTCACCG 1513
|||||
DB 399 ATCTTCGATACAAATGCCAGATATGCATCGTCCTTTCAAGTCCCACTGTTTCACCG 340
|||||

QY 1514 CTTTGTCTTCTACATGCTCTTCATGGTTGCCCTTTCCTCTATTCGACCCATT 1573
|||||
DB 339 CTTTGTCTTCTACATGCTCTTCATGGTTGCCCTTTCCTCTATTCGACCCATT 280
|||||

QY 1574 GTACAGGATTGGCTTCGTCATCATCTGACTGGAGTCCCTCGCTATTATCTTTATTA 1633
|||||
DB 279 GTACAGGATTGGCTTCGTCATCATCTGACTGGAGTCCCTCGCTATTATCTTTATTA 220
|||||

```

QY 1634 TATGGACAAAGAACCCAGGTGGTTTGAATTAATGTCAGAGAAAATACCAAGACATTAC 1693
Db 219 TATGGACAAAGAACCCAGGTGGTTTGAATTAATGTCAGAGAAAATACCAAGACATTAC 160
QY 1694 AAATAACTCTGGAAGTTGTACCAAGAGAGATAAGTATTAAGTAACTGAGATCT 1753
Db 159 AAATAACTCTGGAAGTTGTACCAAGAGAGATAAGTATTAAGTAACTGAGATCT 100
QY 1754 TGGCAATCGCCCAAGGGGAGACACAAAATAGGGATTTTACTTCATTTCTGAAAGTCT 1813
Db 99 TGGCAATCGCCCAAGGGGAGACACAAAATAGGGATTTTACTTCATTTCTGAAAGTCT 40
QY 1814 AGAGAATTACAACCTTTGGTGATAACAAAA 1843
Db 39 AGAGAATTACAACCTTTGGTGATAACAAAA 10

RESULT 3
LOCUS BM887840/c
DEFINITION TMT023 Human Trabecular Meshwork cDNA library Homo sapiens cDNA 5',
mRNA sequence.
ACCESSION BM887840
VERSION BM887840.1 GI:19271584
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 511)
AUTHORS Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
TITLE Expression Profile and Genome Location of cDNA Clones from an
Infant Human Trabecular Meshwork Library
JOURNAL Unpublished (2002)
COMMENT Contact: Wirtz MK
Glaucoma Genetics Lab
Oregon Health Sciences University
3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
Tel: 503-494-4698
Fax: 503-494-6875
Email: wirtzm@ohsu.edu
Seq primer: T7 Reverse.
FEATURES
source
1..511
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human Trabecular Meshwork cDNA library"
/tissue_type="eye"
/cell_type="trabecular meshwork"
/dev_stage="2 week to 2 year old infants"
/lab_host="TOPI0F"
/note="Vector: pCDNA3; Site_1: EcoRI; Site_2: EcoRI; Human
cDNA library made from mRNA isolated from trabecular
meshwork cells established from eyes from 6 individuals,
ages 2 weeks to 2 years. Cells were harvested at passages
3 through 6. Invitrogen made a unidirectional cDNA library
from the mRNA from the frozen cells using a pCDNA3 vector
and TPOI0F" host cells."
BASE COUNT 161 a 99 c 70 g 180 t 1 others
ORIGIN
Query Match 19.6%; Score 438; DB 14; Length 511;
Best Local Similarity 99.8%; Pred. No. 6.5e-120;
Matches 488; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1636 TGGCAAGAACCCAGGTGGTTTGAATTAATGTCAGAGAAAATACCAAGACATTACAA 1695
Db 511 TGGCAAGAACCCAGGTGGTTTGAATTAATGTCAGAGAAAATACCAAGACATTACAA 452
QY 1696 ATAATACTGGAAGTTGTACCAAGAGAGATAAGTATTAAGTAACTGAGATCTTG 1755
Db 451 ATAATACTGGAAGTTGTACCAAGAGAGATAAGTATTAAGTAACTGAGATCTTG 392

```

```

QY 1756 GCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTCATTTCTGAAAGTCTAG 1815
Db 391 GCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTACTTCATTTCTGAAAGTCTAG 332
QY 1816 AGAATTACAACCTTTGGTGATAACAAAAGAGAGTACAGTATTTTATTCATATATTTAGC 1875
Db 331 AGAATTACAACCTTTGGTGATAACAAAAGAGAGTACAGTATTTTATTCATATATTTAGC 272
QY 1876 ATATTGCAACTTAATTTCTAAGAAATTTAGTTATAACTCTATGTAGTTATAGAAAGTGAAT 1935
Db 271 ATATTGCAACTTAATTTCTAAGAAATTTAGTTATAACTCTATGTAGTTATAGAAAGTGAAT 212
QY 1936 ATGCAGTTATTTCTATGAGTCGCACAATTTCTTGAGTCTCTGATACCTACCTATTTGGGGTTA 1995
Db 211 ATGCAGTTATTTCTATGAGTCGCACAATTTCTTGAGTCTCTGATACCTACCTATTTGGGGTTA 152
QY 1996 GGAGAAAAGACTAGACAAATTTACTATGTGGTCATTCTCTACAACATATGTTAGCAGCGCAA 2055
Db 151 GGAGAAAAGACTAGACAAATTTACTATGTGGTCATTCTCTACAACATATGTTAGCAGCGCAA 92
QY 2056 AGAACCTTCAAAATGAAGACTGAGATTTTCTGTATATATATGGGTTTGTAAAGATGGTTT 2115
Db 91 AGAACCTTCAAAATGAAGACTGAGATTTTCTGTATATATATGGGTTTGTAAAGATGGTTT 32
QY 2116 TACACACTA 2124
Db 31 TACACACTA 23

RESULT 4
LOCUS BG490131
DEFINITION 602519050F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4637509 5',
mRNA sequence.
ACCESSION BG490131
VERSION BG490131.1 GI:13451641
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 919)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI398 row: c column: 14
High quality sequence stop: 726.
FEATURES
source
1..919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4637509"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

```

BASE COUNT      253 a   195 c   199 g   272 t
ORIGIN

Query Match      19.2%; Score 429; DB 12; Length 919;
Best Local Similarity 100.0%; Pred. No. 1.8e-117;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1344 GCACCCCTTGCACATGATATGCTCTCTCTGGAGACCTCGACAGCTCTTTGTAATTCCT 1403
Db 237 GCACCCCTTGCACATGATATGCTCTCTCTGGAGACCTCGACAGCTCTTTGTAATTCCT 356
QY 1404 CAGTTTGCAGGTGGCTTTTATTGGCTGGCAGTTGCTGGCTGATTTATCTCGATA 1463
Db 357 CAGTTTGCAGGTGGCTTTTATTGGCTGGCAGTTGCTGGCTGATTTATCTCGATA 416
QY 1464 CAATGCGCCAGATATGATCGTCTCTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTTC 1523
Db 417 CAATGCGCCAGATATGATCGTCTCTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTTC 476
QY 1524 CTTTCACATGCTCTTCATGCTGGCTTTCCTCTATTCGGACCCCAATTTAGTACAGGGAT 1583
Db 477 CTTTCACATGCTCTTCATGCTGGCTTTCCTCTATTCGGACCCCAATTTAGTACAGGGAT 536
QY 1584 TGGCTTGGTCATCACTCTGACTGGAGTCCCTGGCTATTTATCTCTTTATATATGGACAA 1643
Db 537 TGGCTTGGTCATCACTCTGACTGGAGTCCCTGGCTATTTATCTCTTTATATATGGACAA 596
QY 1644 GAAACCCAGGTGGTTAGATAATATGTCAGAGAAAAATACCCAGAACATTTACAAATAATACT 1703
Db 597 GAAACCCAGGTGGTTAGATAATATGTCAGAGAAAAATACCCAGAACATTTACAAATAATACT 656
QY 1704 GGAAGTTGTACCAAGAGAATAAGTTATGAACCTTAATGGACTTGGCAATCTG 1763
Db 657 GGAAGTTGTACCAAGAGAATAAGTTATGAACCTTAATGGACTTGGCAATCTG 716
QY 1764 CCCAAGGGG 1772
Db 717 CCCAAGGGG 725

RESULT 5
LOCUS      BG326527      910 bp      mRNA      linear      EST 27-FEB-2001
DEFINITION 602425373F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562994 5',
            mRNA sequence.
ACCESSION  BG326527
VERSION    BG326527.1 GI:13132964
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE     NIH-MGC http://mgc.nci.nih.gov/
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: DCTD/DTP
          CDNA Library Preparation: Ling Hong/Rubin Laboratory
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LCM1276 row: b column: 19
          High quality sequence stop: 706.
          Location/Qualifiers
            1..910
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:4562994"
              /clone_lib="NIH_MGC_14"

/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: kidney; Vector: pORF7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II-RT (Life Technologies)."
```


/db_xref="taxon:9606"	
/clone="228358"	
/clone_lib="CIT-Hsp"	
/sex="Male"	
/cell_type="Sperm"	
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"	
BASE COUNT	142 a 152 c 133 g 130 t
ORIGIN	
Query Match 7.7%; Score 172; DB 17; Length 557;	
Best Local Similarity 100.0%; Pred. No. 2.2e-41;	
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	335 AGAAGTGCAGCTGAAGAGGAAAGTCACTTACTTACGAGGAGTCTCCATTATCATTTGGCA 394
Db	278 AGAAGTGCAGCTGAAGAGGAAAGTCACTTACTTACGAGGAGTCTCCATTATCATTTGGCA 219
QY	395 CCATCATTTGAGCAGGGAATCTTCATCTCTCTTAAGGCGTGTCCAGAACACGGGCAGCG 454
Db	218 CCATCATTTGAGCAGGGAATCTTCATCTCTCTTAAGGCGTGTCCAGAACACGGGCAGCG 159
QY	455 TGGGCATGCTCTTGACCATCTGACGGGTGTGGGGTCTCTGACATATTTGG 506
Db	158 TGGGCATGCTCTTGACCATCTGACGGGTGTGGGGTCTCTGACATATTTGG 107
RESULT 8	
W00842	
LOCUS	
DEFINITION W00842 371 bp mRNA linear EST 18-APR-1996	
IMAGE:290903 5', mRNA sequence.	
ACCESSION W00842	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
human.	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	
AUTHORS	
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman	
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,	
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston	
R., Williamson,A., Wohldmann,P. and Willson,R.	
The WashU-Merck EST Project	
Unpublished (1995)	
Contact: Wilton RK	
Washington University School of Medicine	
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
Tel: 314 286 1800	
Fax: 314 286 1810	
Email: est@watson.wustl.edu	
This clone is available royalty-free through LLNL; contact the	
IMAGE Consortium (info@image.llnl.gov) for further information.	
Seq primer: Enprimer	
High quality sequence stop: 339.	
Location/Qualifiers	
1. 371	
/organism="Homo sapiens"	
/db_xref="GB:3884033"	
/db_xref="taxon:9606"	
/clone="IMAGE:290903"	
/clone_lib="Soares melanocyte 2NbHM"	
/sex="Male"	
/tissue_type="melanocyte"	
/lab_host="DH10B (ampicillin resistant)"	
/note="Vector: pT73D (Pharmacia) with a modified	
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA	
was primed with a Not I - oligo(dT) primer [5'	
TGTTACCATCTGAGTGGAGCGCCGAGTTTCTTTTCTTTTCTTTT 3'],	
double-stranded cDNA was size selected, ligated to Eco RI	
adapters (Pharmacia), digested with Not I and cloned into	
the Not I and Eco RI sites of a modified pT73 vector	

(Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino.

BASE COUNT 131 a 50 c 69 g 115 t 6 others
ORIGIN

Query Match 7.6%; Score 171; DB 14; Length 371;
Best Local Similarity 99.1%; Pred. No. 6.2e-41;
Matches 321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1661 GAATAATGTCAGAGAAATACCAAGACATTACAATAATCTGGAAGTTGTACCAAGAG 1720
Db 1 GAATAATGTCAGAGAAATACCAAGACATTACAATAATCTGGAAGTTGTACCAAGAG 60

Qy 1721 AAGATAGTTATGAACTTAAGTCTGAGATCTTGGCAATCTGCCAAGGGGAGACACAA 1780
Db 61 AAGATAGTTATGAACTTAAGTCTGAGATCTTGGCAATCTGCCAAGGGGAGACACAA 120

Qy 1781 AATAGGGATTTTACTTCATTTCTGAAAGCTAGAGAATATACAACCTTTTGGTGATAAACA 1840
Db 121 AATAGGGATTTTACTTCATTTCTGAAAGCTAGAGAATATACAACCTTTTGGTGATAAACA 180

Qy 1841 AAAGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAAATTTCTAAGAAAT 1900
Db 181 AAAGAGTCAGTTATTTTATTCATATATTTTAGCATATTCGAACTAAATTTCTAAGAAAT 240

Qy 1901 TTAGTTATACTCTATGTTAGTTATAGAAAGTGAATATGCACTTATTTCTATGAGTCGCACA 1960
Db 241 TTAGTTATATCTATGTTATAGAAAGTGAATATGCACTTATTTCTATGAGTCGCACA 300

Qy 1961 ATTCTTGAGTCTCGATACCTACC 1984
Db 301 ATTCTTGAGTCTCGATACCTACC 324

RESULT 9
A0533681
LOCUS A0533681 400 bp DNA linear GSS 18-MAY-1999
DEFINITION RPCI-11-384K17-TV RPCI-11 Homo sapiens genomic clone RPCI-11-384K17
, DNA sequence.

ACCESSION A0533681
VERSION A0533681.1 GI:4845371
KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 400)
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..400
/organism="Homo sapiens"
/db_xref="GDB:7647328"
/db_xref="taxon:9606"

/clone="RPCI-11-384K17"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector; PBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"
BASE COUNT 103 a 84 c 72 g 140 t 1 others
ORIGIN

Query Match 6.8%; Score 153; DB 17; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.2e-35;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1344 GCACCCCTTTGACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCT 1403
Db 171 GCACCCCTTTGACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCT 230

Qy 1404 CAGTTTGGCAGTGGCTTTTATTTGGGCTGGCAGTCTGGGCTGATTTATCTTCGATA 1463
Db 231 CAGTTTGGCAGTGGCTTTTATTTGGGCTGGCAGTCTGGGCTGATTTATCTTCGATA 290

Qy 1464 CAAATGCCAGATATGATCGCTCTTCAAGGT 1496
Db 291 CAAATGCCAGATATGATCGCTCTTCAAGGT 323

RESULT 10
B33331
LOCUS B33331 486 bp DNA linear GSS 17-OCT-1997
DEFINITION HS-1017-A1-G10-MF.abi CIT Human Genomic Sperm Library C Homo
sapiens genomic clone Plate=CT 792 Col=19 Row=M, DNA sequence.

ACCESSION B33331
VERSION B33331.1 GI:2532700
KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 486)
Mahairas, G.G., Zackrone, K.D., Smith, T., Tipton, S., Schmidt, S., Traicoff, R., Abajian, C., Blanchard, A., West, A. and Hood, L.E.

TITLE Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors
JOURNAL Unpublished (1997)
COMMENT Contact: Mahairas GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kzackrone@u.washington.edu
Sequence Tagged Connector
Plate: CT 792 row: M column: 19
Class: BAC ends
High quality sequence stop: 486.

FEATURES
source
Location/Qualifiers
1..486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=CT 792 Col=19 Row=M"
/clone_lib="CIT Human Genomic Sperm Library C"
/sex="M"
/note="Organ: sperm; Vector: pheloBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 112 a 96 c 109 g 166 t 3 others
ORIGIN

Query Match 6.8%; Score 153; DB 17; Length 486;
Best Local Similarity 100.0%; Pred. No. 1e-35;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1344 GCACCCCTTTGACAATGATAATGCTCTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCT 1403
|||||

Db 62 GCACCCCTTTGACAAATGATATGCTCTCTGGAGACCTCGACAGCTCTTTTGAATTTCT 121
 QY 1404 CAGTTTGGCAGGTGGCTTTTATTTGGCTGCGCAGTCTGCGCTGATTTATCTTCGATA 1463
 Db 122 CAGTTTGGCAGGTGGCTTTTATTTGGCTGCGCAGTCTGCGCTGATTTATCTTCGATA 181
 QY 1464 CAAATGCCAGATATGCATCGCTCTTTCAAGGT 1496
 Db 182 CAAATGCCAGATATGCATCGCTCTTTCAAGGT 214

RESULT 11
 BQ370588
 LOCUS
 DEFINITION PMO-GN0344-271200-001-f12 GN0344 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BQ370588
 VERSION BQ370588.1 GI:21046102
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS 1 (bases 1 to 497)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and
 Simpson, A.J.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.J.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PMO&t2=PMO-GN0344-271200-001-f12&t3=2000-12-27&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 106
 High quality sequence stop: 117.

FEATURES

source
 Location/Qualifiers
 1..497
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="GN0344"
 /dev_stage="Adult"
 /note="Organ: placenta_normal; Vector: puc18; Site: 1: SmaI
 ; Site: 2: SmaI; A mini-library was made by cloning
 products derived from ORESTES PCR (U.S. Letters Patent
 application No. 196,716 - Ludwig Institute for Cancer
 Research) profiles into the puc 18 vector. Reverse
 transcription of tissue mRNA and cDNA amplification were
 performed under low stringency conditions."
 BASE COUNT 103 a 120 c 113 g 161 t
 ORIGIN

Query Match 6.8%; Score 152; DB 14; Length 497;
 Best Local Similarity 100.0%; Pred. No. 2e-35;
 Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1210 TTGGCTCCATGAACGGTGGTGTGTTTGGCTGCTCCAGGTTATCTATGTTGGCTCGA 1269

Db 146 TTGGCTCCATGAACGGTGGTGTGTTTGGCTGCTCCAGGTTATCTATGTTGGCTCGA 205

QY 1270 GAGGGTCACCTTCAGAAATCCCTCCATGATTCATGTCGCGAAGCACACTCTCTACCA 1329

Db 206 GAGGGTCACCTTCAGAAATCCCTCCATGATTCATGTCGCGAAGCACACTCTCTACCA 265
 QY 1330 GCTGTTATCTTTTGGCACCCCTTTGACAATGAT 1361
 Db 266 GCTGTTATCTTTTGGCACCCCTTTGACAATGAT 297

RESULT 12
 BH860963
 LOCUS
 DEFINITION UP_336-3L.SP6 RPi11 Human DNA linear GSS 18-JUL-2002
 clone 336-3L, DNA sequence.

ACCESSION BH860963
 VERSION BH860963.1 GI:21905687
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS 1 (bases 1 to 485)

TITLE Cheung, V.G., Dalrymple, H.L., Narasimhan, S., Watts, J., Schuler, G.,
 Raap, A.K., Morley, M. and Bruzel, A.
 JOURNAL A resource of mapped human bacterial artificial chromosome clones
 MEDLINE Genome Res. 9 (10), 989-993 (1999)
 COMMENT 99455100

Contact: Narasimhan SL, Morley M, Burdick J, Cheung VG
 Department of Pediatrics
 University of Pennsylvania
 3516 Civic Center Blvd, ARC 516, Philadelphia, PA 19104, USA
 Tel: 215 590 2664
 Fax: 215 590 3709
 Email: sandyaz@email.med.upenn.edu
 Plate: 336 row: L column: 3
 Seq primer: SP6
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..485
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="336-3L"
 /clone_lib="RPi11 Human Male BAC Library"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBACe3.6; RPi11 Human Male BAC Library"
 BASE COUNT 132 a 97 c 93 g 163 t
 ORIGIN

Query Match 5.7%; Score 128; DB 17; Length 485;
 Best Local Similarity 100.0%; Pred. No. 2.5e-28;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 749 CTGTAGTGTGCTTCTTAATAGCATGAGTGTGAGTGGAGCGCCGGATCCAGATTTTCT 808

Db 249 CTGTAGTGTGCTTCTTAATAGCATGAGTGTGAGTGGAGCGCCGGATCCAGATTTTCT 308

QY 809 TAACCTTTTGCAGAGCTCACAGCAATCTGATATATATAGTCCCTGGAGTTATGACAGTAA 868

Db 309 TAACCTTTTGCAGAGCTCACAGCAATCTGATATATATAGTCCCTGGAGTTATGACAGTAA 368

QY 869 TTAAGAGT 876

Db 369 TTAAGAGT 376

RESULT 13
 BG388107
 LOCUS
 DEFINITION 602413070f1 NTH_MGC_92 Homo sapiens cDNA clone IMAGE:4521736 5',
 mRNA sequence.

ACCESSION BG388107
 VERSION BG388107.1 GI:13281553
 KEYWORDS EST.

SOURCE
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1072)
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue procurement: ARCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM0421 row: k column: 17
 High quality sequence stop: 625.
FEATURES
 source
 1..1072
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4521736"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
BASE COUNT 331 a 229 c 270 g 242 t
ORIGIN

Query Match 5.0%; Score 112; DB 12; Length 1072;
 Best Local Similarity 100.0%; Pred. No. 6.7e-24;
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 391 GGCACCATCTGGAGCAGGAATCTTCATCTCTCTAAGCGCGTCTCCAGAACACGGC 450
 |||||||
 Db 452 GGCACCATCTGGAGCAGGAATCTTCATCTCTAAGCGCGTCTCCAGAACACGGC 511
 |||||||

Qy 451 AGCGTGGCGATGCTCTGACCATCTGGACGGTGTGGGTGCTGTGCATCAT 502
 |||||||
 Db 512 AGCGTGGCGATGCTCTGACCATCTGGACGGTGTGGGTGCTGTGCATCAT 563
 |||||||

RESULT 14
LOCUS BG186884 752 bp mRNA linear EST 21-APR-2001
DEFINITION RST5861 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG186884
VERSION BG186884.1 GI:13708571
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
 Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,
 Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith,
 E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher,
 J., Danzig, J. and Ducar, M.
TITLE Creation of genome-wide protein expression libraries using random
 activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave., Cleveland, OH 44115, USA
 Tel: 216 431 9900

Fax: 216 361 9596
 Email: scain@atersys.com
 High quality sequence stop: 441.
FEATURES
 source
 1..752
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Athersys RAGE Library"
 /cell_line="HT1080"
 /note="See 'Creation of Genome-wide protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is HT1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in HT1080 under normal circumstances."
BASE COUNT 246 a 147 c 132 g 225 t
ORIGIN

Query Match 4.6%; Score 102; DB 12; Length 752;
 Best Local Similarity 100.0%; Pred. No. 8.2e-21;
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1983 CCTATTGGGTTAGGAGAAAGACTAGACAATTACTATGTGTCATTCTTACACATAT 2042
 |||||||
 Db 278 CCTATTGGGTTAGGAGAAAGACTAGACAATTACTATGTGTCATTCTTACACATAT 219
 |||||||

Qy 2043 GTTAGCAGCGCAAGAACCTTCAAAATTGAAGACTGAGATTTT 2084
 |||||||
 Db 218 GTTAGCAGCGCAAGAACCTTCAAAATTGAAGACTGAGATTTT 177
 |||||||

RESULT 15
LOCUS AW059922/c 87 bp mRNA linear EST 23-AUG-2000
DEFINITION LF1b04.Yg UPCI5 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW059922
VERSION AW059922.1 GI:6652244
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 87)
 Brenner, S., Williams, S.R., Vermaas, E.H., Storch, T., Moon, K.,
 McCollum, C., Mao, J.I., Kirchner, J.J., Eletre, S., Dubridge, R.B.,
 Burcham, T. and Albrecht, G.
TITLE In vitro cloning of complex mixtures of DNA on microbeads: Physical
 separation of differentially expressed cDNAs
 Proc. Natl. Acad. Sci. U.S.A. 97 (4), 1665-1670 (2000)
JOURNAL 20144098
MEDLINE Contact: Burcham TS
COMMENT LYNX Therapeutics, Inc.
 25861 Industrial Blvd., Hayward, CA 94545, USA
 Tel: 510 670 9338
 Fax: 510 670 9302
 Email: timbelynxgen.com
 Sequence obtained from LYNX Therapeutics Megasort technology.
 Collected from the up-regulated gate.
 High quality sequence stop: 87.
FEATURES
 source
 1..87
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="UPC15"
 /cell_type="monocytic leukemia"
 /cell_line="THP-1 (TIB-202)"
 /note="Vector: pCR2.1; Cloning of PCR products from
 micro-beads carrying 3' end of up-regulated cDNA. THP-1
 cells induced with 100 nM PMA in DMSO."
BASE COUNT 27 a 19 c 14 g 27 t
ORIGIN

Query Match 3.9%; Score 87; DB 10; Length 87;

Best Local Similarity 100.0%; Pred. No. 1.4e-15;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1749 GATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAATTTCTGAA 1808
 Db 87 GATCTTGGCAATCTGCCCAAGGGGAGACACAAAAATAGGGATTTTACTTCAATTTCTGAA 28
 QY 1809 AGTCTAGAGAATTACAACCTTTGGTGAT 1835
 Db 27 AGTCTAGAGAATTACAACCTTTGGTGAT 1

Search completed: April 16, 2003, 14:10:03
 Job time : 3558 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 06:13:54 ; Search time 93 Seconds
(without alignments)
7383.325 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 2239

Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	366.4	16.4	2072	US-09-073-362-2	Sequence 2, Appli
2	366.4	16.4	2072	US-09-243-920-2	Sequence 2, Appli
3	211.4	9.4	1011	US-08-825-781-2	Sequence 2, Appli
4	67.4	3.0	7218	US-08-232-463-14	Sequence 14, Appl
5	51.6	2.3	62804	US-09-800-960-3	Sequence 3, Appli
6	50	2.2	152331	US-09-128-155-16	Sequence 16, Appli
7	44.8	2.0	6124	US-08-213-419B-3	Sequence 3, Appli
8	43.6	1.9	5375	US-08-757-223-7	Sequence 7, Appli
9	43.4	1.9	893	US-09-370-838-166	Sequence 166, App
10	43.4	1.9	40000	US-09-780-049-18	Sequence 18, Appl
11	43	1.9	99500	US-09-798-096-10	Sequence 10, Appl
12	43	1.9	162450	US-09-345-882-1	Sequence 1, Appli
13	42.8	1.9	70000	US-09-851-896-3	Sequence 3, Appli
14	42.8	1.9	72604	US-09-268-992-7	Sequence 7, Appli
15	42.8	1.9	72604	US-09-657-474-7	Sequence 7, Appli
16	42.6	1.9	3224	US-08-965-729A-2	Sequence 2, Appli
17	42.6	1.9	12047	US-09-022-461-1	Sequence 1, Appli
18	42.6	1.9	12047	US-09-033-556-3	Sequence 3, Appli
19	42.4	1.9	72604	US-09-268-992-7	Sequence 7, Appli
20	42.4	1.9	72604	US-09-657-474-7	Sequence 7, Appli
21	42.2	1.9	1825	US-09-461-697-75	Sequence 75, Appl
22	42.2	1.9	7218	US-08-232-463-14	Sequence 14, Appl
23	41.8	1.9	301	US-08-332-766A-23	Sequence 23, Appl
24	41.8	1.9	22481	US-08-367-841A-43	Sequence 43, Appli
25	41.8	1.9	22481	PCT-US95-07201-43	Sequence 2, Appli
26	41.8	1.9	22484	US-09-875-223-2	Sequence 2, Appli
27	41.6	1.9	55627	US-09-813-133A-3	Sequence 3, Appli

28 41.4 1.8 471 4 US-09-018-584A-6 Sequence 6, Appli
c 29 41.4 1.8 81001 4 US-09-750-580-1 Sequence 1, Appli
c 30 41.2 1.8 516 4 US-09-018-584A-24 Sequence 24, Appl
c 31 41 1.8 6405 4 US-09-281-481A-18 Sequence 18, Appl
32 41 1.8 17327 1 US-07-906-871-15 Sequence 15, Appl
33 41 1.8 38564 4 US-09-734-673-3 Sequence 3, Appli
c 34 41 1.8 111282 4 US-09-754-250-3 Sequence 3, Appli
35 40.8 1.8 87350 3 US-08-781-891-79 Sequence 79, Appl
36 40.8 1.8 87543 4 US-09-791-211-3 Sequence 3, Appli
37 40.6 1.8 45546 4 US-09-146-053-6 Sequence 6, Appli
38 40.4 1.8 2821 2 US-08-680-395-6 Sequence 6, Appli
39 40.2 1.8 4820 4 US-08-961-527-19 Sequence 19, Appl
c 40 40.2 1.8 12597 4 US-09-705-299-12 Sequence 12, Appl
41 40.2 1.8 14581 4 US-08-520-373D-4 Sequence 4, Appli
c 42 40.2 1.8 28720 4 US-09-341-587-7 Sequence 7, Appli
43 40 1.8 1001 4 US-09-641-638-401 Sequence 401, App
c 44 40 1.8 8453 4 US-09-167-681-45 Sequence 45, Appli
45 39.8 1.8 162450 4 US-09-345-882-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-073-362-2
; Sequence 2, Application US/09073362
; Patent No. 5942399
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ceirone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0514 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2072 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: ESOGTUT02
; CLONE: 2667831
; US-09-073-362-2

Query Match 16.4% Score 366.4; DB 2; Length 2072;
Best Local Similarity 54.6%; Pred. No. 2.2e-92;

Qy	300	CGCCCTTTTCAGGAAGAGACGCCCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAAGT	359
Db	192	CTCCCTTTTGGGTGATGTGGGGCCAGGCCAGGGCGGAGCAGGTGAAGCTGAAGAAGGAGAT	251
Qy	360	CACTTTACTCAGGGGAGTCTCCATTATCATTTGSCACCATCATTTGAGCAGCAAGTCTTCAT	419
Db	252	CTCACTGCTTTAACGGGCTGTGCCCTGATGTGGGGAACATGATCGGCTCAGGCATCTTTGT	311
Qy	420	CTCTCTAAGGGCTGCTCTCCAGAACACGGCAGCGTGGGCATGTCTTGACCATCTGGAC	479
Db	312	TTCCCCCAAGGGTGTGCTCATATACAGTGCCTCCCTTTGGTCTCTCTCTGGTCACTGGC	371
Qy	480	GGTGTGTGGGTCCTGTCACTATTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAT	539
Db	372	TGTGGGGGGCTCTCTCCGTCTTTGGGGCCCTTTGTTATGCGGAACCTGGGCACCCAT	431
Qy	540	AAAGAAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTTGGTCCATTACCAAGCTTT	599
Db	432	TAAGAAATCTGGGCCAGCTATGCCCTATATCTGAGGCCCTTTGGAGGATTCCTTGTCTTT	491
Qy	600	TGTACGAGTCTGGGTGGAACCTCCTCATAATAGCCCTGCGAGCTACTGCTGTGATATCCCT	659
Db	492	CATCAGACTCTGGACCTCCCTGCTCATTCATTGAGCCACCAGCCAGGCCATCAITGGCAT	551
Qy	660	GGCATTGGAGCGTACATTCCTGGAACCATTTTTTATTCAAATGGAATTCCTCGAATTCG	719
Db	552	CACCTTTTGGCAACTACATGGTACAGCCTCTCTTCCGAGCTGCTTTCGCCCTTATGTGTC	611
Qy	720	GATCAAGCTCATTTACAGCTGTGGGCATAACTCTGATGATGGTCTTAATAGCATGAGTGT	779
Db	612	CAGCGGCTGCTGGTGTGCTGCTGCATTTGTCTCTTAACCTTCATTAAGTGTGCTGATGT	671
Qy	780	CAGCTGAGGCGCCGATCCAGATTTTCTTAACCTTTTGCAGCTCACAGCAATTCCTGAT	839
Db	672	CAATGGGAACCCCTGGTACAAGATATTTTCAACCTATGCTAAAGTATTGGCAGCTATCGC	731
Qy	840	AATTAATAGTCCCTGGAGTTATGCAGCTAATTTAAAGGTCAAAGCGAAGATTTAAGAGCG	899
Db	732	GGTCATCGTTTGCAGGCAATTTGTAGACTTGGCCAGGAGGCCCTCTACTCATTTTGGAAATC	791
Qy	900	CTTTTTCAGGAAGAGATTTCAAGTATTACGGGTGGCCACTGGCTTTTATATATGAAGTGA	959
Db	792	CTTTGAGGGTTCATCATTTGCAGTGGTGACATTTGCCCTGGCAGCTGTACTCAGCTCTGTT	851
Qy	960	TGCATATGCTGGTGGTGTTFACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCTGAAAA	1019
Db	852	CTCTACTCAGGCTGGGACACCCCTCAACTATGTCATGAAGAGATCAAGAATCTTGAGAG	911
Qy	1020	AACATTCGCCCTTGCATATGATATATCCATGGCCATTTGTCACCATTGGCTATGTGCTGAC	1079
Db	912	GAACTCGCCCTCTCCATTTGGCATCTCCATGGCCATTTGTCACCACTCATATATCTTGAC	971
Qy	1080	AAATGTGCCCTACTTTACAGACCATTAATGCTGAGGAGCTGTGCTTTCCAAATCAGATGGC	1139
Db	972	CAATGTGCCCTATTATACTGTCTGACATGAGAGACATCTTGGCCAGTATGCTGTGTC	1031
Qy	1140	AGTGACCTTTTTCGAGGGGTACTGGGAAATTTCTCATTTAGCAGTTCGATCTTTGTGTGC	1199
Db	1032	TGTGACTTTTTCAGATCAGATATTTTGGAAATATTTAACTGGATATTTCCACTGTCTGATGC	1091
Qy	1200	CCTCTCTGCTTTGGCTCCATGAAGGTGGTGTGTTTGTCTGTCTCCAGGTTATTCATGT	1259
Db	1092	ATTATCTGTTTGGTGGCTCTAATGCCCTCAATGGCTGTGCTCTTAGGCTTTTCTTTGT	1151
Qy	1260	TGGCTCTCGAGAGGGTCACCTTCCAGAAATCCTCTCCATGATTTCATGTCGCAAGCACAC	1319
Db	1152	GGGCTCAAGAAGAGGCCATCTCCCTGATGCCATCTGCATGATCCATGTTGAGCGTTCAC	1211
Qy	1320	TCCTCTACCAAGCTGTTATTTGTTTTGACCCCTTTGACAAATGATAATGCTCTCTCGGAGA	1379
Db	1212	ACAGTGCCTCTCTGCTCTCAATGGTATCATGCGATTCGATCTACTGTGCGGTGGAAGA	1271

Qy	1380	CCTCGACAGCTCTTTTGAATTCCTCAGTTTTTCCAGGTGGCTTTTTATTGGGCTGGCAGT	1439
Db	1272	CATCTTCCAGCTCATTAACACTACTACAGCTTCAGCTACTTGGTTCCTTTGGGGGCTTCTCAT	1331
Qy	1440	TGCTGGGCTCATTTATCTTCGTACAAATGCCAGATATGCATCGTCTCTTTCAAGGTGCC	1499
Db	1332	TGTGGGTCAGCTTATCTGGCTGGNAGGAGCCTGATCGACCTGTCCCTCAAGCTCAG	1391
Qy	1500	ACTGTTATATCCAGCTTTGTTTCTTCATATGCCTCTTCATATGGCCTTTGCCCTTCCCTCTTA	1559
Db	1392	CGTTTCTTCCCGAATTGTCTTCTGGCTCTGTCACCATCTTCTGGTGGCTGTTCACACTTTA	1451
Qy	1560	TTCGGACCATTTAGTACAGGGATTGGCTTGGCTATCACTCTGACTGGAGTCCCTCGCTA	1619
Db	1452	CAGTGATACTATCAACTCCCTCATCGGCAATGCCCTCTCAGGCCCTGCCCTTTTA	1511
Qy	1620	TTATCTCTTTATTATA	1635
Db	1512	CTTCCCTCATCATCAGA	1527

RESULT 3

```

US-08-825-781-2
: Sequence 2, Application US/08825781
: Patent No. 5843727
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Shah, Purvi
: TITLE OF INVENTION: HUMAN TUMOR-ASSOCIATED MEMBRANE PROTEIN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/825,781
: FILING DATE: Herewith
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0262 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1011 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: PANCTUT02
: CLONE: 2236771
US-08-825-781-2

```

Query Match 9.4%; Score 211.4; DB 2; Length 1011;
Best Local Similarity 56.3%; Pred. No. 2.8e-49;
Matches 395; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

Db 21 CCTGGCAGTCTACTAGCTCTGTCTCTCTACTAGGCTGGACACCTCAACTATGTCA 80
QY 995 CTGAAGAAGTAGAAGCCCTCAAAAACCAATTCCTCTTGGCAATATGTATATCCATGCCA 1054
Db 81 CTGAAGAAGTAGAAGCCCTCAAAAACCAATTCCTCTTGGCAATATGTATATCCATGCCA 140
QY 1055 TTGTCACCAATGGTATGTCTGACAAATGTGGCTACTTTACGACCAATTAATGCTGAGG 1114
Db 141 TTGTCACCAATGGTATGTCTGACAAATGTGGCTACTTTACGACCAATTAATGCTGAGG 200
QY 1115 AGCTGCTCTCCAGGTTATCTATGTTGGCTCGAGAGGCTCACTTCCAGAAATCCTCT 1294
Db 201 ACATCTTGGCAGTGTCTGCTGTGACCTTTTGCAGATCAGATATTTGGAATATTTA 260
QY 1175 CATTAGCAGTCCGATCTTGTGCTCCCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 1234
Db 261 ACTGGATAATTCCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320
QY 1235 TTGCTGTCTCCAGGTTATCTATGTTGGCTCGAGAGGCTCACTTCCAGAAATCCTCT 1294
Db 321 TGGCTGCTTCTAGGCTTTCTTCTTGGGCTCAAGAGAAGGCCATCTCCCTGATGCCATCT 380
QY 1295 CCATGATTCATGTCGGAAGCACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1354
Db 381 GCATGATTCATGTCGGAAGCACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 440
QY 1355 CAATGATAATGCTCTCTCTGAGACCTCGACAGCTCTTTTGAATTTCTCTAGTTTGGCA 1414
Db 441 CATTGATCTACTTGGCTGGAAGACATCTTCCAGCTCACTTAATCTACTACAGCTTCAGCT 500
QY 1415 GGTGGCTTTTATTTGGCTGGCAGTCTGCTGGCTGATTTATCTTCGATACAAATGCCAG 1474
Db 501 ACTGGTCTTGTGGGCTTTCTATTTGGCTGAGCTTATCTGCTGCTGCTGCTGCTGCTGCTG 560
QY 1475 ATATGATCTGCTTCAAGTGCCACTGTTCATCCAGCTTTTCTCTCTCTCTCTCTCTCTCT 1534
Db 561 ATGACACTCTGCTCCCTCAAGCTCAGCTGCTTTCTTCCGCTGCTCTCTCTCTCTCTCTCT 620
QY 1535 TCTTCATGTTGGCT 1594
Db 621 TCTTCTGTTGGCTGCT 680
QY 1595 TCATCTGACTGGAGTCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1635
Db 681 TTGCTCTGAGGCTGCT 721

RESULT 4

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14
Query Match 3.0%; Score 67.4; DB 1; Length 7218;
Best Local Similarity 2.1%; Pred. No. 1.3e-08;
Matches 8; Conservative 235; Mismatches 136; Indels 0; Gaps 0;
QY 1184 TTCGATCTTTGTTGCCCT 1243
Db 1060 TTGCGATTTTCTATGTTGGCTCTCGAGAGGCTCACTTCCAGAAATCCTCTCCATGATTC 1119
QY 1244 CCAGGTTTATCTATGTTGGCTCTCGAGAGGCTCACTTCCAGAAATCCTCTCCATGATTC 1303
Db 1120 YY 1179
QY 1304 ATGTCGCAAGCACACCT 1363
Db 1180 YY 1239
QY 1364 TGCTCTCTCTGAGAGCTCGACAGCTCTTTGAATTTCCCTCAGTTTTCGCCAGTGCTTTT 1423
Db 1240 YY 1299
QY 1424 TTATTTGGCTGGCAGTTGCTGGCTGATTTATCTTCGATACAAATGCCAGATATGATC 1483
Db 1300 YY 1359
QY 1484 GTCCTTTCAAGGTGCCACTGTTTATCCAGCTTTTGTCTTCCCTCACATGCTCTTCATGG 1543
Db 1360 YY 1419
QY 1544 TTGCTCTTCT 1562
Db 1420 YYYYYYYYYYYYYYYYGTAC 1438
RESULT 5
US-09-800-960-3/c
; Sequence 3, Application US/09800960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001158
; CURRENT APPLICATION NUMBER: US/09/800,960
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804

Query Match 1.9%; Score 43.6; DB 3; Length 5375;
Best Local Similarity 67.8%; Pred. No. 0.049;
Matches 61; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 GGAGGTGAAGTGGAGAGAGATCATCCAGCCTGGGTGACAGTGGAGACTCTGTCTCAAAAC 60
DB 2550 GGGAGTTACAGTGGAGCAACATTGTCCAGCCTGGGTGACAGAGAGAGACTCTGTCTCAA 2609
QY 61 AGAATTAAGGAAAAAGAAAAAGAAAAA 90
DB 2610 AAAAAAAGAAAAAAGAAAAAAGAAAAA 2639

RESULT 9

US-09-370-838-166
; Sequence 166, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 166
; LENGTH: 893
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-166

Query Match 1.9%; Score 43.4; DB 4; Length 893;
Best Local Similarity 62.4%; Pred. No. 0.021;
Matches 68; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 14 AGCAGAGATCATCCAGCCTGGGTGACAGTGGAGACTCTGTCTCAAAACAGAAATTAAGGAAA 73
DB 88 AACTGAGGCTCGGTCGGAAGGGTGAGTGGAGACTACATCTCAAAAAAAGAAAAA 147
QY 74 AAGAAAAAGAAAAAGAGAGAGAGAAATTCAGGCCAAATGTGGCA 122
DB 148 AAAAAAGAAAAAGAAAAAGAAAAAAGAAAAAGAACGGAAGTAGTTGTA 196

RESULT 10

US-09-780-049-18/c
; Sequence 18, Application US/09780049
; Patent No. 6465250
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PROTEIN PHOSPHATASE 2 CATALYTIC SUBUNIT
; FILE REFERENCE: RTS-0134
; CURRENT APPLICATION NUMBER: US/09/780,049
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 40000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-049-18

Query Match 1.9%; Score 43.4; DB 4; Length 40000;
Best Local Similarity 69.4%; Pred. No. 0.17; Indels 26; Mismatches 0; Gaps 0;

QY 23 CATGCCAGCCTGGGTGACAGTGGAGACTCTGTCTCAAAACAGAAATTAAGGAAAAAAGA 82
DB 7124 CACTCCAGCCTGGGAGCTGTGAGACTCTATCTCAAAAAAAGAAAAAAGAAAAA 7065
QY 83 AAGAAAAAGAGAGAGAGAAATTC 107
DB 7064 AAGCCACAAAAACAAAAATCTTC 7040

RESULT 11

US-09-798-096-10/c
; Sequence 10, Application US/09798096
; Patent No. 6399378
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF RECOL2 EXPRESSION
; FILE REFERENCE: RTS-0207
; CURRENT APPLICATION NUMBER: US/09/798,096
; CURRENT FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 10
; LENGTH: 99500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-798-096-10

Query Match 1.9%; Score 43; DB 4; Length 99500;
Best Local Similarity 77.6%; Pred. No. 0.36;
Matches 52; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 23 CATGCCAGCCTGGGTGACAGTGGAGACTCTGTCTCAAAACAGAAATTAAGGAAAAAAGA 82
DB 32580 CACTCCAGCCTGGGTGACAGAGCGAGACTGTCTCAAAAAAAGAAAAAAGAAAAA 32521

QY 83 AAGAAAA 89
DB 32520 CGGAAGA 32514

RESULT 12

US-09-345-882-1/c
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele

US-09-864-761-48172

Alignment Scores:

Pred. No.:	142	Length:	34
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	10	Gaps:	0

US-09-667-170A-440 (1-2239) x US-09-864-761-48172 (1-34)

QY 70 GAAAAAGAAAGAAAGAAAGAG 93

Db 24 GluLysArgLysLysGluLysGlu 31

Search completed: April 16, 2003, 16:57:06
Job time : 48 secs

Alignment Scores: 142 Length: 34
Pred. No.: 8.00 Matches: 8
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.09% Gaps: 0
DB: 10
US-09-667-170A-440 (1-2239) x US-09-864-761-41510 (1-34)
QY 70 GAAAGAGAGAGAGAGAGAGAGAG 93
DB 24 GlnLysArgLysGluLysGlu 31
RESULT 15
US-09-864-761-48172
; Sequence 48172, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48172
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALL18523.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9

QY 1993 TTAGGAGAGAGAGACTAGACAACTACTA 2019
DB 44 LeuGlyGluLysThrArgLysLeu 52
RESULT 14
US-09-864-761-41510
; Sequence 41510, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41510
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALL133419.11
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
US-09-864-761-41510

Pred. No.: 1.19 Length: 523
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.37% Indels: 0
DB: 9 Gaps: 0

US-09-667-170A-440 (1-2239) x US-10-163-866-47 (1-523)

QY 409 GGAATCTTCATCTCTCTAAGGCGCTGCTC 438

Db 57 GlyIlePheIleSerProIyglyValIeu 66

RESULT 11

US-09-739-907-99

; Sequence 99, Application US/09739907

; Patent No. US20010012889A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 36 Human Secreted Proteins

; FILE REFERENCE: P2022P1

; CURRENT APPLICATION NUMBER: US/09/739,907

; CURRENT FILING DATE: 2000-12-20

; PRIOR APPLICATION NUMBER: 09/348,457

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: 60/070,567

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,692

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,704

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,658

; PRIOR FILING DATE: 1998-01-07

; NUMBER OF SEQ ID NOS: 196

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 99

; LENGTH: 132

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (132)

; OTHER INFORMATION: Xaa equals stop translation

US-09-739-907-99

Alignment Scores:

Pred. No.: 13 Length: 132
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 10 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-739-907-99 (1-132)

QY 1993 TTAGGAGAAAGACTAGACAATTACTA 2019

Db 34 LeuGlyIleSerProIyglyValIeu 42

RESULT 12

US-09-739-907-87

; Sequence 87, Application US/09739907

; Patent No. US20010012889A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 36 Human Secreted Proteins

; FILE REFERENCE: P2022P1

; CURRENT APPLICATION NUMBER: US/09/739,907

; CURRENT FILING DATE: 2000-12-20

; PRIOR APPLICATION NUMBER: 09/348,457

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: 60/070,567

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (170)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-739-907-87

Alignment Scores:

Pred. No.: 12.4 Length: 172
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 10 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-739-907-87 (1-172)

QY 1993 TTAGGAGAAAGACTAGACAATTACTA 2019

Db 34 LeuGlyIleSerProIyglyValIeu 42

RESULT 13

US-09-739-907-191

; Sequence 191, Application US/09739907

; Patent No. US20010012889A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 36 Human Secreted Proteins

; FILE REFERENCE: P2022P1

; CURRENT APPLICATION NUMBER: US/09/739,907

; CURRENT FILING DATE: 2000-12-20

; PRIOR APPLICATION NUMBER: 09/348,457

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: 60/070,567

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,692

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,704

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,658

; NUMBER OF SEQ ID NOS: 196

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 191

; LENGTH: 182

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (180)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-739-907-191

Alignment Scores:

Pred. No.: 12.3 Length: 182
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 10 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-739-907-191 (1-182)

Alignment Scores:		
Pred. No.:	0.016	511
Scores:	12.00	12
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	1.64%	0
DE:	9	0
Length:		
Matches:		12
Conservative:		0
Mismatches:		0
Indels:		0
Gaps:		0

Alignment Scores:

RESULT 4

US-09-738-973-586
 ; Sequence 586, Application US/09738973
 ; Patent No. US20020110563A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Fling, Steven P.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Indirias, Carol Yoseph
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Elliot, Mark
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Kalos, Michael D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; THE THERAPY AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.475C9
 ; CURRENT APPLICATION NUMBER: US/09/738,973
 ; CURRENT FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 587
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 586
 ; LENGTH: 97
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-738-973-586

Alignment Scores:

Pred. No.: 3,17e-78 Length: 97
 Score: 93.00 Matches: 93
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 12.70% Indels: 0
 DB: 10 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-738-973-586 (1-97)

QY 2 GAGTTGAAGTGGAGCAGATCATGCCAGCCTGGGTGACAGTGCAGACTCTCTCAACA 61
 |||||
 Db 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20
 QY 62 GAATTAAGCAAAAAGAAAGAAAAGAGAGAGAGAGAAATCCAGGCCAATTTGGC 121
 |||||
 Db 21 GluLeuArgLysLysGluArgLysLysArgGluArgLysPheGlnAlaAsnCysGly 40
 QY 122 ATAGATTTTATCATATCTGGATTTTGGATTTCTTTGTTTCTCATCTGATTCAG 181
 |||||
 Db 41 IleAspPheIlePheThrPheThrPheThrPheThrPheThrPheThrPheThrPhe 60
 QY 182 GAAACCTGTGTGCCACCATCTCCAAAGAGGTTACCTGCAGGGAAATGTTAACGGGA 241
 |||||
 Db 61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80
 QY 242 GCGTCCTCTCCCTGGCAACAGAGCCACCTGGCCAGG 280
 |||||
 Db 81 GlyCysLeuProTrrPAlaThrArgSerHisLeuGlyArg 93

RESULT 5

US-09-854-133-587
 ; Sequence 587, Application US/09854133
 ; Publication No. US20020183499A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Secrist, Heather
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; THE THERAPY AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.475C10
 ; CURRENT APPLICATION NUMBER: US/09/854,133
 ; CURRENT FILING DATE: 2001-05-11
 ; NUMBER OF SEQ ID NOS: 735
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 587
 ; LENGTH: 16
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-854-133-587

Alignment Scores:

Pred. No.: 5,23e-06 Length: 16
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.19% Indels: 0
 DB: 9 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-854-133-587 (1-16)

QY 104 TTCCAGGCCAATTTGGCATAGATTTTATCATATTTCTGGATTTTGG 151
 |||||
 Db 1 PheGlnAlaAsnCysGlyLeaSpPheIleIlePheTrpIlePheTrp 16

RESULT 6

US-09-738-973-587
 ; Sequence 587, Application US/09738973
 ; Patent No. US20020110563A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Fling, Steven P.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Indirias, Carol Yoseph
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Elliot, Mark
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Kalos, Michael D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; THE THERAPY AND DIAGNOSIS OF LUNG CANCER

Alignment Scores:

Pred. No.: 5,23e-06 Length: 16
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.19% Indels: 0
 DB: 10 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-738-973-587 (1-16)

QY 104 TTCCAGGCCAATTTGGCATAGATTTTATCATATTTCTGGATTTTGG 151
 |||||
 Db 1 PheGlnAlaAsnCysGlyLeaSpPheIleIlePheTrpIlePheTrp 16

RESULT 7

US-10-163-866-41
 ; Sequence 41, Application US/10163866
 ; Publication No. US20030027188A1

Db 36 GluLysValGlnLeuLysArgLysValThrLeuLeuArgGlyValSerIleIleIleGly 55
QY 394 ACCATCATTTGGAGCAGGAATCTTCTCTCTAAGGCGGTGCTCCAGAACACGGGACG 453
Db 56 ThrIleIleGlyValGlyIlePheIleSerProLysGlyValLeuGlnAsnThrGlySer 75
QY 454 GTGGCATGCTCTCTGACCATCTGGAGCGGTGTGTGGGTCTCTACTATTGGAGCTTTG 513
Db 76 ValGlyMetSerLeuThrIleIleIleIleIleIleIleIleIleIleIleIleIle 95
QY 514 TCTTATGCTGAATTTGGGAACACTATAAGAAATCTGGAGGTCATTACACATATATTGT 573
Db 96 SerTyrAlaGluLeuGlyThrThrIleLysLysSerGlyGlyHisTyrThrIleLeu 115
QY 574 GAAGTCTTTGGTCCATTACCATGCTTTGTACGAGTCTGGGTGGAATCTCTCATATAACG 633
Db 116 GluValPheGlyProLeuProAlaPheValArgValTyrValGluLeuIleIleArg 135
QY 634 CTGTCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTGGAAACATTTT 693
Db 136 ProAlaAlaThrAlaValIleSerLeuAlaPheGlyArgTyrIleLeuGluProPhePhe 155
QY 694 ATTCAATGTGAATCCCTGAATCTGGATCAAGCTCATTTACAGCTGTGGGCATAACTGTA 753
Db 156 IleGlnCysGluLeuProGluLeuAlaIleLysLeuIleThrAlaValGlyIleThrVal 175
QY 754 GTGATGCTCTTAATAGCATGTGTCAGTGGAGCGCCGGATCCAGATTTCTTAACC 813
Db 176 ValMetValLeuAsnSerMetSerValSerTrpSerAlaArgIleGlnIlePheLeuThr 195
QY 814 TTTTGCAGCTCACCAATCTCGATAATATAGTCCCTGGAGTTATGAGCTCAATATAA 873
Db 196 PheCysLysLeuThrAlaIleLeuIleIleIleIleIleIleIleIleIleIleIle 215
QY 874 GGTCAACCGCAGAACTTTAAGACGCTTTTTCAGGAAGATTCAGTATTACGCGGTG 933
Db 216 GlyGlnThrGlnAsnPheLysAspAlaPheSerGlyArgAspSerIleThrArgLeu 235
QY 934 CCAGTGGCTTTTATGAAATGTATGATATGCTGCTGGCTTTTACCTCAACTTTGT 993
Db 236 ProLeuAlaPheTyrGlyMetTyrAlaTyrAlaGlyTrpPheTyrLeuAsnPheVal 255
QY 994 ACTGAAGAAGTAGAAAACCTGAAAACCATTCCTCCCTTGCAATATGATATCATGGCC 1053
Db 256 ThrGluGluValGluAsnProGluLysThrIleProLeuAlaIleCysIleSerMetAla 275
QY 1054 ATTCACCAATGCTGCTATGCTGACAAATGTGGCTACTTTACGACCATTAATGCTGAG 1113
Db 276 IleValThrIleGlyTyrValLeuThrAsnValAlaTyrPheThrIleAsnAlaGlu 295
QY 1114 GAGCTGCTGCTTTCAAAATGAGTGGCAGTGCACCTTTTCTGAGCGGCTACTGGGAAATTC 1173
Db 296 GluLeuLeuLeuSerAsnAlaValAlaValThrPheSerGluArgLeuLeuGlyAsnPhe 315
QY 1174 TCATTAGCAGTCCGATCTTTGTTGGCCTCTCCCTGCTTGGCTCCATGACGGTGGTG 1233
Db 316 SerLeuAlaValProIlePheValAlaLeuSerCysPheGlySerMetAsnGlyVal 335
QY 1234 TTTGCTGCTCCAGTTATCTATGTTGCTGCTGAGGCTGCTCCAGAAATCCTC 1293
Db 336 PheAlaValSerArgLeuPheTyrValAlaSerArgGluGlyHisLeuProGluIleLeu 355
QY 1294 TCCATGATTATGTCGCGACACACTCTCTACAGCTGTATTGTTTTCACCCCTTG 1353
Db 356 SerMetIleHisValArgLysHisThrProLeuProAlaValIleValLeuHisProLeu 375
QY 1354 ACAATGATAATGCTCTCTGAGACCTCGACAGTCTTTTGAATTTCCCTCAGTTTGGC 1413
Db 376 ThrMetIleMetLeuPheSerGlyAspLeuAspSerLeuLeuAsnPheLeuSerPheAla 395
QY 1414 AGTGGCTTTTATTGGCTGGCAGTCTGCTGGCTGATTTATCTTCGATACAAATGCCCA 1473
Db 396 ArgTrpLeuPheIleGlyLeuAlaValAlaGlyLeuIleTyrLeuArgTyrLysCysPro 415

QY 1474 GATATGCATCGTCTCTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTCTTCCATCATGC 1533
Db 416 AspMetHisArgProPheLysValProLeuPheIleProAlaLeuPheSerPheThrCys 435
QY 1534 CTCTTCATGTTGCCCTTTCCCTCTATTCGACCCCAATTTAGTACAGGATTTGGCTTCGTC 1593
Db 436 LeuPheMetValAlaLeuSerLeuTyrSerAspProPheSerThrGlyIleGlyPheVal 455
QY 1594 ATCACTCTGACTGAGTCCCTGCTGATATCTCTTTATATATATGGGACAAAGAACCCAGG 1653
Db 456 IleThrLeuThrGlyValProAlaTyrTyrLeuPheIleIleIleTrpAspLysLysProArg 475
QY 1654 TGGTTTGAATAATGTCCAGAAATAACAGAACATTTACAAATAATPACTGGAGTTGTA 1713
Db 476 TrpPheArgIleMetSerGluLysIleThrArgThrLeuGlnIleIleLeuGluValVal 495
QY 1714 CCAGAAGAAGATAAGTTA 1731
Db 496 ProGluGluAspLysLeu 501
RESULT 3
US-09-854-133-586
; Sequence 586, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854.133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 586
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-133-586
Alignment Scores:
Pred. No.: 3.17e-78 Length: 97
Score: 93.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.70% Indels: 0
DB: 9 Gaps: 0
US-09-667-170A-440 (1-2239) x US-09-854-133-586 (1-97)
QY 2 GAGGTGAAGTCAGCAGAGATCATGCCCTGGGTGACAGTGCAGTGCAGTCTGTCTCAACA 61
Db 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20
QY 62 GAATTAAGGAAAAAGAAAAAGAAAAAGAGAGAGAGAGAAATTCAGGCCAATTTGTGGC 121
Db 21 GluLeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGly 40
QY 122 ATAGATTTTATCATATCTGGATTTTTCGATTTTCTTCTTCTCATCTGGATTTCAG 181
Db 41 IleAspPheIleIlePheIlePheIlePheIlePheIlePheIlePheIlePheIle 60
QY 182 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGGTTCCTCGAGGGAATGTTACAGGGA 241
Db 61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80
QY 242 GGCTGCTTCCCTGGGCACACAGGACCCACCTGGGCGAG 280
Db 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArg 93

Alignment Scores:

Pred. NO.:	0	Length:	501
Score:	466.00	Matches:	466
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	63.66%	Indels:	0
DB:	9	Gaps:	0

US-09-667-170A-440 (1-2239) x US-10-163-866-48 (1-501)

334	QY	GAGAAAGTCGACCTGAAGAGAAAGTCACCTTTACTGAGGGAGTCTCCATATTCATTGGC	339	DB	GLuLysValGLnLeuLysArgLysValThrLeuLeuArgGLyValSerIleIleIleGLy	344	QY	ACCATCATTTGGAGCAGGAATCTTCATCTCTCCCTAAAGGGCGTGCTCCAGACACGGGAGC	349	DB	ThrIleIleGLyAlaGLyIlePheIleSerProLysGLyValLeuGLnAsnThrGLySer	354	QY	GTGGGCGATGCTCTGACCATCTGGACCGGTGTGTGGGGCTCTGTCACTATTTTGGAGCTTGG	359	DB	ValGLyMetSerLeuThrIleTprThrValCysGLyValLeuSerLeuPheGLyAlaLeu	364	QY	TCCTTATGCTGAATTTGGGAACAACATATAAGAAATCTGGAGTCAATTACACATATATTTTG	369	DB	SerTYrAlaGLuLeuGLyThrIleLysLysSerGLyHisTYrThrTYrIleLeu	374	QY	GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGAACCTCCTCATATAATACGC	379	DB	GLuValPheGLyProLeuProAlaPheValArgValIrrpValGLuLeuLeuIleIleArg	384	QY	CCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTCTGGAACCATTTTGT	389	DB	ProAlaIleThrAlaValIleSerLeuAlaPheGLyArgTYrIleLeuGLuProPhePhe	394	QY	ATTCAATGCTGAATTCCTCGAACTTCGGCATCAAGCTCATACAGCTGTGGGCATACTGTA	399	DB	IleGLnCysGLuIleProGLuLeuAlaIleLysLeuIleThrAlaValGLyIleThrVal	404	QY	GTGATGTGCTTAATAGCATGTGACTGTACAGCTGGAGCGCCGGATCCAGATTTCTTAACC	409	DB	ValMetValLeuAsnSerMetSerValSerIrrpSerAlaArgIleGLnIlePheLeuThr	414	QY	TTTTGCCAAGCTCACAGCAATTCGTGAATAATATAGTCCCTGGAGTTATGCAGCTAATAAA	419	DB	PheCysLysLeuThrAlaIleLeuIleIleValProGLyValMetGLnLeuIleLys	424	QY	GGTCAAAACGCAGAACCTTTAAACAGCCCTTTTCAGGACAGATTCAGTATTACGGCGTTG	429	DB	GlyGLnThrGLnAsnPhelyAspAlaPheSerGLyArgAspSerIleThrArgLeu	434	QY	CCACTGGCTTTTATTATGGAATGTATGCATATGCTGGCTGGTTTTCACCTCAACTTGT	439	DB	ProLeuAlaPheTYrTYrGLyMetTYrAlaTYrAlaGLyIrrpPheTYrLeuAsnPheVal	444	QY	ACTCAAGAAGTAGAAACCCCTGAAAACCAATCCCTTGCAATATGTATATCCATGGCC	449	DB	ThrGLuGLuValGLuAsnProGLuLysThrIleProLeuAlaIleCysIleSerMetAla	454	QY	ATTGTCACCAATTGGCTATGTGCTGCAGAAATGTGGCTACTTTACGACCAATTAACTGTGAG	459	DB	IleValThrIleGLyTYrValLeuThrAsnValAlaTYrPheThrThrIleAsnAlaGLu	464	QY	GAGTGTGCTTTTCAATTCAGTGGCAGTGCACCTTTTCTTGAGCGGCTACTGGGAAATTC	469	DB	GLuLeuLeuLeuSerAsnAlaValAlaValThrPheSerGLuArgLeuGLyAsnPhe	474	QY	TCATTAGCAGTTCCCATCTTTGTGGCCCTCTCTGCTTTGGCTCCATGAACGGTGTGTG	479	DB	SerLeuAlaValProIlePheValAlaLeuSerCysPheGLySerMetAsnGLyGLyVal	484	QY	TTTGCTGTCTCCAGGTTAATCTATGTGTGGCTCTCGAGAGGGTCACTTCGAGAATCCCT	489	DB	PheAlaValSerArgLeuPheTYrValAlaSerArgGLuGLyHisLeuProGLuIleLeu	494	QY		499	DB		504	QY		509	DB		514	QY		519	DB		524	QY		529	DB		534	QY		539	DB		544	QY		549	DB		554	QY		559	DB		564	QY		569	DB		574	QY		579	DB		584	QY		589	DB		594	QY		599	DB		604	QY		609	DB		614	QY		619	DB		624	QY		629	DB		634	QY		639	DB		644	QY		649	DB		654	QY		659	DB		664	QY		669	DB		674	QY		679	DB		684	QY		689	DB		694	QY		699	DB		704	QY		709	DB		714	QY		719	DB		724	QY		729	DB		734	QY		739	DB		744	QY		749	DB		754	QY		759	DB		764	QY		769	DB		774	QY		779	DB		784	QY		789	DB		794	QY		799	DB		804	QY		809	DB		814	QY		819	DB		824	QY		829	DB		834	QY		839	DB		844	QY		849	DB		854	QY		859	DB		864	QY		869	DB		874	QY		879	DB		884	QY		889	DB		894	QY		899	DB		904	QY		909	DB		914	QY		919	DB		924	QY		929	DB		934	QY		939	DB		944	QY		949	DB		954	QY		959	DB		964	QY		969	DB		974	QY		979	DB		984	QY		989	DB		994	QY		999	DB	
-----	----	---	-----	----	--	-----	----	---	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	---	-----	----	---	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	---	-----	----	---	-----	----	---	-----	----	--	-----	----	--	-----	----	--	-----	----	---	-----	----	---	-----	----	--	-----	----	--	-----	----	--	-----	----	---	-----	----	---	-----	----	--	-----	----	--	-----	----	---	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--	-----	----	--

QY	1294	TC	CATGATTCATGTC	CGCAAGCACAC	CTCTCTAC	CAGCTGTTATG	TTTGTGCAC	CTTTG	1353														
Db	356	Ser	Met	Ile	Ile	His	Val	Arg	Gly	His	Thr	Pro	Leu	Pro	Ala	Val	Ile	Val	Leu	His	Pro	Leu	375
QY	1354	ACA	AGTAA	TAA	AGCTCT	TCTCTG	GAGACCTCG	ACAGCTCTTT	GAA	TTTCTC	AGTTT	GGCC	1413										
Db	376	Thr	Met	Ile	Met	Leu	Phe	Ser	Gly	Asp	Leu	Asp	Ser	Leu	Leu	Asn	Phe	Leu	Ser	Phe	Ala	395	
QY	1414	AGG	TGGCTTT	TAT	TGGCTGG	CAGAGT	GTGGCTG	CTATTAT	CTTC	TGCATAC	AAATG	CCCA	1473										
Db	396	Arg	Tp	Leu	Phe	Ile	Gly	Leu	Ala	Val	Ala	Gly	Leu	Ile	Tyr	Leu	Arg	Tyr	Ile	Cys	Pro	415	
QY	1474	GAT	TGCAT	CTCTCT	TCA	AGGTGC	CACCTG	TCTATC	CCAGCTTT	GTTTTC	CTC	CATGC	1533										
Db	416	Asp	Met	His	Arg	Pro	Phe	Lys	Val	Pro	Leu	Phe	Ile	Pro	Ala	Leu	Phe	Ser	Phe	Thr	Cys	435	
QY	1534	CTC	TTCA	TGTTG	CGCCCTT	CCCTCTAT	TCGGAC	CCATTTAG	TAC	AGGATTTG	CGCTCG	TC	1593										
Db	436	Leu	Phe	Met	Val	Ala	Leu	Ser	Leu	Tyr	Ser	Asp	Pro	Phe	Ser	Thr	Gly	Ile	Gly	Phe	Val	455	
QY	1594	ATC	ACTCTG	ACTGG	AGTCC	TGGTATTAT	TC	TCTTTATAT	TGG	CACAA	GAACCC	AGG	1653										
Db	456	Ile	Thr	Leu	Thr	Gly	Val	Pro	Ala	Tyr	Tyr	Leu	Phe	Ile	Ile	Thr	Asp	Lys	Lys	Pro	Arg	475	
QY	1654	TGG	TTT	AGA	TATAT	GC	CAG	AAAAAT	AACC	CAC	ACATTTAC	AAAT	TACTG	G	AA	TTG	T	G	T	A	1713		
Db	476	Trp	Phe	Arg	Ile	Met	Ser	Glu	Lys	Ile	Thr	Arg	Thr	Leu	Gln	Ile	Ile	Leu	Glu	Val	Val	495	
QY	1714	CC	AGA	GAAG	AT	AAG	T	A	1731														
Db	496	Pro	Glu	Glu	Asp	Lys	Leu	501															

RESULT 2

US-10-163-866-49

; Sequence 49, Application US/10163866

; Sequence 457, Application 03/101
; Publication No. US20030027188A1

GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

FILE REFERENCE: EX02-080C

; CURRENT APPLICATION NUMBER: US/1

; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: 2001-06-05

; PRIOR APPLICATION NUMBER: US
 : PRIOR FILING DATE: 2001-06-05

; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: US 60/4228 007

; PRIOR APPLICATION NUMBER: US
:
: PRIOR FILING DATE: 2001-10-10

; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/338 733

; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-10-22

; PRIORITY FILING DATE: 2001-10-22
 ; PRIORITY APPLICATION NUMBER: US 60/357 353

; PRIORITY AFFILIATION NUMBER: US
; PRIORITY FILING DATE: 2002-02-15

PRIOR APPLICATION NUMBER: US 60/357,600

PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 54

```

; SOFTWARE: PatentIn version 3.1

```

; SEQ ID NO 49

; LENGTH: 501

```

; TYPE: PRT

```

; ORGANISM: Homo sapiens

US-10-163-866-49

Alignment Scores:

Argument scores:					
Pred. No.:	0	Length:	501		
Score:	466.00	Matches:	466		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	93.66%	Indels:	0		
DB:	6	Gaps:	0		

US-09-667-170A-440 (1-2239) x US-10-163-866-49 (1-501)

QY 334 .GAGAAAGTGCAGCTGAAGAGGAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC 393

Thu Apr 17 09:48:31 2003

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Alignment Scores:
Pred. No.:      16.6      Length:      2227
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.12%      Indels:      0
DB:              3          Gaps:      0

US-09-667-170A-440 (1-2239) x US-08-475-886-6 (1-2227)

QY 990 AAGTTGAGGTAAACCCAGC 967
|||||
Db 1350 LysValGluValProAlaSer 1357

RESULT 15
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Alignment Scores:
Pred. No.:      16.6      Length:      2227
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.12%      Indels:      0
DB:              4          Gaps:      0

US-09-667-170A-440 (1-2239) x US-08-397-232-2 (1-2227)

QY 990 AAGTTGAGGTAAACCCAGC 967
|||||
Db 1350 LysValGluValProAlaSer 1357

Search completed: April 16, 2003, 16:42:51
Job time : 33.5 secs

```

```

; TITLE OF INVENTION: AND VACCINES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04910
; FILING DATE: 21-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US08/231,526
; FILING DATE: 22-APR-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4125PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 995 amino acids
; TYPE: amino acids
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; PCT-US95-04910-14

Alignment Scores:
Pred. No.: 18, 3 Length: 995
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 5 Gaps: 0

US-09-667-170A-440 (1-2239) x PCT-US95-04910-14 (1-995)
QY 1727 TTATCTTCTCTGGTCAACTCC 1704
Db 434 LeuSerSerSerGlyThrThrSer 441

RESULT 12
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT

US-09-667-170A-440 (1-2239) x US-08-475-886-2 (1-2227)
QY 990 AAAGTTGAGTAAACACGCCAGC 967
Db 1350 LysValGluValLysProAlaSer 1357

RESULT 13
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT

US-08-475-886-4
Alignment Scores:
Pred. No.: 16, 6 Length: 2227
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 3 Gaps: 0

US-09-667-170A-440 (1-2239) x US-08-475-886-4 (1-2227)
QY 990 AAAGTTGAGTAAACACGCCAGC 967
Db 1350 LysValGluValLysProAlaSer 1357

RESULT 14
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT

```

```

; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Alignment Scores:
Pred. No.: 16, 6 Length: 2227
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 3 Gaps: 0

US-09-667-170A-440 (1-2239) x US-08-475-886-2 (1-2227)
QY 990 AAAGTTGAGTAAACACGCCAGC 967
Db 1350 LysValGluValLysProAlaSer 1357

RESULT 13
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT

US-08-475-886-4
Alignment Scores:
Pred. No.: 16, 6 Length: 2227
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 3 Gaps: 0

US-09-667-170A-440 (1-2239) x US-08-475-886-4 (1-2227)
QY 990 AAAGTTGAGTAAACACGCCAGC 967
Db 1350 LysValGluValLysProAlaSer 1357

RESULT 14
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT

```

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 4 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-556-273-11 (1-786)

QY 766 AATAGCATGAGTGTGCAGCTGGAGC 789

Db 549 AsnSerMetSerValSerTrpSer 556

RESULT 9

US-09-012-710-10

; Sequence 10, Application US/09012710

; Patent No. 6087478

; GENERAL INFORMATION:

; APPLICANT: Vinkemeier, Uwe

; APPLICANT: Moarefi, Ismail

; APPLICANT: Darnell, Jr., James E.

; APPLICANT: Kuriyan, John

; TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson

; STREET: 411 Hackensack Avenue, 4th Floor

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/012.710

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-194

; TELEPHONE: 201-487-5800

; TELEFAX: 201-343-1684

; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 793 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

US-09-012-710-10

Alignment Scores:

Pred. No.: 18.8 Length: 793

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.09% Indels: 0

DB: 3 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-012-710-10 (1-793)

QY 766 AATAGCATGAGTGTGCAGCTGGAGC 789

Db 549 AsnSerMetSerValSerTrpSer 556

RESULT 10

US-09-556-273-10

; Sequence 10, Application US/09556273
; Patent No. 6312887
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; APPLICANT: Moarefi, Ismail
; APPLICANT: Darnell, Jr., James E.
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/556,273

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/012,710

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.

; REGISTRATION NUMBER: 26,742

; REFERENCE/DOCKET NUMBER: 600-1-194

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-487-5800

; TELEFAX: 201-343-1684

; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 793 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

US-09-556-273-10

Alignment Scores:

Pred. No.: 18.8 Length: 793

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.09% Indels: 0

DB: 4 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-556-273-10 (1-793)

QY 766 AATAGCATGAGTGTGCAGCTGGAGC 789

Db 549 AsnSerMetSerValSerTrpSer 556

RESULT 11

PCT-US95-04910-14

; Sequence 14, Application PC/TUS9504910

; GENERAL INFORMATION:

; APPLICANT: The Government of the United

; APPLICANT: States of America as represented

; APPLICANT: by the Secretary, Department of

; APPLICANT: Health and Human Services

; TITLE OF INVENTION: ISOLATION AND

; TITLE OF INVENTION: CHARACTERIZATION OF A NOVEL PRIMATE T-CELL

; TITLE OF INVENTION: LYMPHOTROPIC VIRUS AND THE USE OF THIS VIRUS

; TITLE OF INVENTION: OR COMPONENTS THEREOF IN DIAGNOSTIC ASSAYS

ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0514 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 515 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1665759
US-09-243-920-3

Alignment Scores:
Pred. No.: 19.8 Length: 515
Score: 8.00 Matches: 8
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
DB: 0

US-09-667-170A-440 (1-2239) x US-09-243-920-3 (1-515)

QY 517 TATGCTGAATTGGGAACAATATA 540
|||||
Db 98 TyralaGlulLeuGlyThrThrle 105

RESULT 7
US-09-012-710-11
Sequence 11, Application US/09012710
Patent No. 6087478
GENERAL INFORMATION:
APPLICANT: Vinkemeier, Uwe
APPLICANT: Moarefi, Ismail
APPLICANT: Darnell, Jr., James E.
APPLICANT: Kuriyan, John
TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
TITLE OF INVENTION: STAT PROTEIN AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,710
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/012,710
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-194
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 786 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-556-273-11

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-012-710-11

Alignment Scores:
Pred. No.: 18.8 Length: 786
Score: 8.00 Matches: 8
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
DB: 0

US-09-667-170A-440 (1-2239) x US-09-012-710-11 (1-786)

QY 766 AATAGCATGAGTCAGCTGGAGC 789
|||||
Db 549 AsnSerMetSerValSerTrpSer 556

RESULT 8
US-09-556-273-11
Sequence 11, Application US/09556273
Patent No. 6312887
GENERAL INFORMATION:
APPLICANT: Vinkemeier, Uwe
APPLICANT: Moarefi, Ismail
APPLICANT: Darnell, Jr., James E.
APPLICANT: Kuriyan, John
TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
TITLE OF INVENTION: STAT PROTEIN AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/556,273
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/012,710
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-194
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 786 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-556-273-11

Alignment Scores:
Pred. No.: 18.8 Length: 786
Score: 8.00 Matches: 8

US-09-071-035-306
; Sequence 306, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 306:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-306

Alignment Scores:
Pred. No.: 21.3 Length: 275
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 4 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-071-035-306 (1-275)

QY 437 AGCAGCCCTTAGGAGAGATGAAG 414
Db 125 SerThrProLeuGlyGlutMetlys 132

RESULT 5

US-09-073-362-3
; Sequence 3, Application US/09073362
; Patent No. 5942399
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrione, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0514 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1665759
US-09-073-362-3
Alignment Scores:
Pred. No.: 19.8 Length: 515
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0
US-09-667-170A-440 (1-2239) x US-09-073-362-3 (1-515)
QY 517 TATGCTGAATGGGAACAATATA 540
Db 98 TyrAlaGluLeuGlyThrThrIle 105
RESULT 6
US-09-243-920-3
; Sequence 3, Application US/09243920
; Patent No. 5981242
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,920
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/073,362

TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 511 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: ESOQTUT02
; CLONE: 2667831
US-09-073-362-1

Alignment Scores:
Pred. No.: 0.00219 Length: 511
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-073-362-1 (1-511)

Qy 517 TATCTGAATGGGAACAACATAAAGAATCTGGA 552
Db 90 TyrAlaGluLeuGlyThrThrIleLysSergly 101

RESULT 2

US-243-920-1
; Sequence 1, Application US/09243920
; Patent No. 5981242
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,920
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/073,362
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ceitron, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0514 US
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 511 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: ESOQTUT02
; CLONE: 2667831
US-09-243-920-1

Alignment Scores:
Pred. No.: 0.00219 Length: 511
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 2 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-243-920-1 (1-511)

Qy 517 TATCTGAATGGGAACAACATAAAGAATCTGGA 552
Db 90 TyrAlaGluLeuGlyThrThrIleLysSergly 101

RESULT 3

US-09-071-035-308
; Sequence 308, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 308:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-308

Alignment Scores:
Pred. No.: 21.6 Length: 241
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 4 Gaps: 0

US-09-667-170A-440 (1-2239) x US-09-071-035-308 (1-241)

Qy 437 AGCAGCCCTTAGAGAGATGAAG 414
Db 96 SerThrProLeuGlyGluMetLys 103

RESULT 4

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 16, 2003, 16:33:19 ; Search time 18.5 Seconds

(without alignments)
7121.938 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 732

Sequence: 1 ggaagttgaagtgcagag.....ttattataaaaaaaaaa 2239

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09667170/runat_04042003_091002_21223/app_query.fasta_1.2375
-DB=Issued Patents_AA -QFMT=fastan -SUFFIX=olin2p.ra -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09667170 -CGN_1_17 @runat_04042003_091002_21223 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCAPOP=6 -FCAPEXT=7
-YCAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/6C.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	12	1.6	511	2	US-09-073-362-1
2	12	1.6	511	2	Sequence 1, Appli
3	8	1.1	241	4	US-09-243-920-1
4	8	1.1	241	4	Sequence 1, Appli
5	8	1.1	275	4	US-09-071-035-308
6	8	1.1	515	2	Sequence 306, App
7	8	1.1	515	2	Sequence 3, Appli
8	8	1.1	786	3	US-09-243-920-3
9	8	1.1	786	3	Sequence 3, Appli
10	8	1.1	793	3	US-09-012-710-11
11	8	1.1	793	3	Sequence 11, Appl
12	8	1.1	995	5	US-09-556-273-11
					Sequence 11, Appl
					Sequence 10, Appl
					Sequence 10, Appl
					Sequence 14, Appl
					Sequence 2, Appli

c 13	8	1.1	2227	3	US-08-475-886-4	Sequence 4, Appli
c 14	8	1.1	2227	3	US-08-475-886-6	Sequence 6, Appli
c 15	8	1.1	2227	4	US-08-397-232-2	Sequence 2, Appli
c 16	8	1.1	2227	4	US-08-397-232-4	Sequence 4, Appli
c 17	8	1.1	2227	4	US-09-171-387-2	Sequence 2, Appli
c 18	8	1.1	2227	4	US-09-653-499-2	Sequence 2, Appli
c 19	8	1.1	2227	4	US-09-653-499-4	Sequence 2, Appli
c 20	8	1.1	2227	4	US-09-653-499-6	Sequence 2, Appli
c 21	7	1.0	12	2	US-08-406-330-44	Sequence 44, Appli
c 22	7	1.0	12	2	US-08-556-597-44	Sequence 44, Appli
c 23	7	1.0	18	1	US-07-972-032-8	Sequence 8, Appli
c 24	7	1.0	18	1	US-08-642-255-8	Sequence 8, Appli
c 25	7	1.0	21	4	US-08-817-548A-3	Sequence 3, Appli
c 26	7	1.0	22	3	US-08-940-095-156	Sequence 156, App
c 27	7	1.0	22	3	US-08-940-095-158	Sequence 158, App
c 28	7	1.0	22	3	US-08-940-095-166	Sequence 166, App
c 29	7	1.0	22	3	US-08-940-095-172	Sequence 172, App
c 30	7	1.0	22	3	US-08-940-095-177	Sequence 177, App
c 31	7	1.0	22	3	US-08-940-095-184	Sequence 184, App
c 32	7	1.0	22	3	US-08-940-093-156	Sequence 156, App
c 33	7	1.0	22	3	US-08-940-093-158	Sequence 158, App
c 34	7	1.0	22	3	US-08-940-093-166	Sequence 166, App
c 35	7	1.0	22	3	US-08-940-093-172	Sequence 172, App
c 36	7	1.0	22	3	US-08-940-093-177	Sequence 177, App
c 37	7	1.0	22	3	US-08-940-093-184	Sequence 184, App
c 38	7	1.0	22	3	US-08-940-096-156	Sequence 156, App
c 39	7	1.0	22	3	US-08-940-096-158	Sequence 158, App
c 40	7	1.0	22	3	US-08-940-096-166	Sequence 166, App
c 41	7	1.0	22	3	US-08-940-096-172	Sequence 172, App
c 42	7	1.0	22	3	US-08-940-096-177	Sequence 177, App
c 43	7	1.0	22	3	US-08-940-096-184	Sequence 184, App
c 44	7	1.0	22	4	US-09-465-719-156	Sequence 156, App
c 45	7	1.0	22	4	US-09-465-719-158	Sequence 158, App

ALIGNMENTS

RESULT 1

US-09-073-362-1
; Sequence 1, Application US/09073362
; Patent No. 5942399
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: AMINO ACID PERMEASE HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PP-0514 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166


```

Db 1726 TGGAT 1730
|||||
RESULT 14
AB011263
LOCUS Homo sapiens mRNA for MOP-2, complete cds. PRI 15-AUG-2000
DEFINITION AB011263
ACCESSION AB011263
VERSION AB011263.1 GI:9836571
KEYWORDS MOP-2.
SOURCE Homo sapiens monocyte mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1536)
Takayama,K. and Yoshimoto,M.
Molecular and Biological Characterization of a Novel Monocyte Amino
Acid Permease, MOP-2
Unpublished
2 (bases 1 to 1536)
Takayama,K. and Yoshimoto,M.
Direct Submission
Submitted (23-FEB-1998) Kiyoshi Takayama, Taisho Pharmaceutical Co.,
Ltd., Molecular Biology Laboratory; Yoshino-cho 1-403, Ohmiya,
Saitama 330, Japan (E-mail:sl1677@cm.taisho.co.jp,
Tel: +81-48-663-1111, Fax: +81-48-652-7254)
FEATURES
source
1..1536
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="monocyte"
1..1536
/notes="similar to a amino acid permease"
/codon_start=1
/product="MOP-2"
/protein_id="BAB11849.1"
/db_xref="GI:9836572"
/translation="MVDSTYEYVASQPEVETSLPDGASPGPEQVKLKKIEILLNGVC
LIVNMGIFGIFSPKGLVILSVIAWGLFVFGALCVAEGLTIIKKSGA
SYAILIEAFGQFLAFLRLTSLIIIEPTSOAILAITFANYVQPLFPSCFAPYASRL
LAAACILITFINCAVVKWGLVDQIFITAKVLALIANIVAGIVLQGQASHFENSF
EGSFVAGDIALALYSALFSYGDNLNIVTEIKNPERNLPLSGISMPVITIIIL
TNVAYITLMDRLASDAVATFDQIFENWIIPLSVALSCEFGNLNLSIVASRL
FFVSGREGLPDALICMHWRETPVPSLLFNGLIMALIYLCVEDIFQLINYSFYSWFF
VGLSIVGQLYLRWKEPRPLKLSVFFIFCLCTIFLVAVPLXSDTINSLIGLIAIA
LSGLPFEVLIIRVPEHKRPLYLRRIVGSATRYLVQLCMSVAEMDLEDGEMPKORDP
KSN"
BASE COUNT 303 a 420 c 364 g 449 t
ORIGIN
Query Match 16.4%; Score 366.4; DB 9; Length 1536;
Best Local Similarity 54.6%; Pred. No. 9.5e-76;
Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;
QY 300 CGCCTTTTCAGGAAGACAGCCCTTTTCAGGAAGAGAGAAAGTCAGCTGAAGAGAAAGT 359
DB 51 CTCCTCTTGGGTGATGGGGCCAGCCCGAGGCGGAGAGGTGAAGCTGAAGAAGAGAT 110
QY 360 CACTTTTACTGAGGGAGCTCTCCATATATTCATTTGGCCACCATCATTTGGAGCAGGATCTTCAT 419
DB 111 CTCACTGCTTACGGCGGTGCTGATTTGGGGAACATATCGGCTCAGGCATCTTTGT 170
QY 420 CTCTCTTAAAGGGGTGCTCCAGAACAGGGCAGGCGAGGTCCTCTCTGACCACATCTGGAC 479
DB 171 TTCCCCCAAGGGGTGCTCATATACAGTGCCCTCTCTTGGTCTCTCTGTCATCTGGGC 230
QY 480 GGTGTGGGGTCTCTGACATATTTGGAGCTTTGTCTTATGCTGAATGGGAACAACATAT 539
DB 231 TGTGGGGGGCTCTCTCTCCCTCTCTGGGGCCCTTTGTTATGCGAACTGGGGCACCACCAT 290
QY 540 AAGAAATCTGGAGGTCAATACATATATTTTGGAAAGCTTTGGTCCATACCACTTT 599
|||||

```

```

Db 291 TAAGAAATCTGGGGCCAGCATATGCCTATATCTCTGGAGGCTTTGGAGGATTCTCTGCTTT 350
QY 600 TGTACAGAGTCTGGGTGAAACTCTCTATAATAGCCCTCTGAGCTACTGCTGTGATATPCCT 659
DB 351 CATCAGACTCTGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 410
QY 660 GGCATTCTGGAGCGTACATCTCTGGAACCAATTTTATTTCAATGTGAATCCCTGAACTTGC 719
DB 411 CACCTTTGCCAAGCTACATCTGTCAGAGCTCTCTCTCCGAGCTCTCTGCGCCCTTATGCTGC 470
QY 720 GATCAAGCTCATTTACAGCTGTGGGCATAACTGTAGTGTAGTGGTCTTAAATAGCATGAGTGT 779
DB 471 CAGCCGCTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 530
QY 780 CAGCTGGAGGGCGGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGTAT 839
DB 531 CAAATGGGAAACCTGGTCAAGATATTTTCACTATGCTAAAGTATTGGCACTGATCGC 590
QY 840 AATTATAGTCCCTGGAGTTATGCAGCTAATTAAGAGTCAACGCGAAGCACTTTAAAGACGC 899
DB 591 GGTCACTGTTGCAGGCAATTTAGACTTGGCAGGAGGCTCTACTCATTTTGAAGATTC 650
QY 900 CTTTTTCAGGAAGAGATCAAGTATTTACGGGTGGCCACTGGCTTTTATATATGGAATGTA 959
DB 651 CTTTGGAGGTTCATCTATTTTGCAGTGGGTGACATTTGCCCTGGCAGCTGCTACTCAGCTCTGT 710
QY 960 TGCATATGCTGGCTGGTCTTACTCTCAACTTTGTTACTGAAGAGTGAAGAAACCTTGAANA 1019
DB 711 CTCTACTCAGGCTGGGACACCTCAACTATGTCACTGAAGAGATCAAGAAATCCCTGGAGAG 770
QY 1020 AACCATTTCCCTTGCATATATGATATCCATGGCCATTTGACCATTTGGCTGTGCTGAC 1079
DB 771 GAACCTGGCCCTCTCCATTTGGCATCTCCATGCCATTTGTCACCATCATCTATCTTGAC 830
QY 1080 AATGTGGCTCTACTTTACGACCATTAATCTGAGGAGCTGCTGCTTTTAAATGTCAGTGGC 1139
DB 831 CAATGTGGCTATATATATCTGTCTAGCATGAGAGACATCTTGGCAGTGTGCTGTTC 890
QY 1140 AGTGACCTTTTCTGAGCGCTACTGGGAATTTCTCATTTAGCAGTTCCGATCTTGTGTC 1199
DB 891 TGTGACTTTTGCAGATCAGATATTTTGAATATTTTAACTGGATAATTTCCACTGTCAGTTGC 950
QY 1200 CCTCTCTCTTGGCTCCATGAAGCGGTGGTGTGTTGCTGCTCCAGGTTATTTCTATGT 1259
DB 951 ATTATCTCTTTTGGTGGCTCAATGCCCTCCATTTGGCTGCTCTAGGCTTTTCTTGT 1010
QY 1260 TGGCTCTCAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCGCAAGACAC 1319
DB 1011 GGGCTCAAGAGAAGGCCCATCTCCCTGATGCCATCTGCATGATCCATGTTGAGCGGTTCCAC 1070
QY 1320 TCTCTTACCAGCTGTATTTGTTTGGCACCTTTTGACAATGATAATGCTCTCTCTGAGAGA 1379
DB 1071 ACCAGTGCCTTCTGCTCTTCAATGGTATCATGGCATTTGATCTACTTGTGGTGGAGA 1130
QY 1380 COTCGACAGCTTTTGAATTTTCCCTCAGTTTGGCCAGGTGGCTTTTATTTGGCTGGCAGT 1439
DB 1131 CATCTTCCAGCTCATTAACCTACTACAGCTTACAGTGTGTTCTTTTGTGGGCTTTCTAT 1190
QY 1440 TGTGGGTGATTTATCTTCGATACAAATGCGCCAGATATGCAATGCTCTCTTCAAGGTGCC 1499
DB 1191 TGTGGGTGAGTCTTATCTGCGCTGGAAGAGGCTGTGACACCTGCTCCCTCAAGCTCAG 1250
QY 1500 ACTGTTTCATCCAGCTTTTGTCTTCTTCTCAGATGCTCTTCAATGCTTGGCTTCCCTCTTA 1559
DB 1251 CGTTTTCTTCCGATTTGCTTCTGCTCTGCAACCATCTTCTCTGGTGGCTGTTTCCACTTTA 1310
QY 1560 TTCGACCACTTTTAGTACAGGATTTGGCTTCGTCTCATCTACTCTGAGTGAGTCCCTTGCCTA 1619
DB 1311 CAGTGATATACTCAACTCCCTCATCGGCATTTGCCATTTGCCCTCTCAGGCTGCCCTTTTA 1370
QY 1620 TTATCTCTTTATATA 1635
DB 1371 CTTCCTCATCATCAGA 1386

```

```

RESULT 13
AF019906          2263 bp      mRNA      linear      VRT 16-MAR-1999
LOCUS              Xenopus laevis L amino acid transporter-1 LAT-1 mRNA, complete cds.
DEFINITION
ACCESSION          AF019906
VERSION            AF019906.1  GI:2731773
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 2263)
Torres, D., Estevez, R., Pineda, M., Fernandez, E., Lloberas, J.,
Shi, Y.B., Zorzano, A. and Palacin, M.
Identification and characterization of a membrane protein (y+L
amino acid transporter-1) that associates with 4F2hc to encode the
amino acid transport activity y+L. A candidate gene for lysinuric
protein intolerance
J. Biol. Chem. 273 (49), 32437-32445 (1998)
JOURNAL
MEDLINE            99047611
PUBMED             9829974
REFERENCE          2 (bases 1 to 2263)
AUTHORS            Shi, Y.
Direct Submission
Submitted (18-AUG-1997) Lab. of Mol. Embryology, National Institute
of Child, Health and Human Development, NIH, Building 18T, Rm 106,
Bethesda, MD 20892, USA
Location/Qualifiers
1..2263
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IU12"
/clone_lib="Patterton, et al., Dev Biol 167, 252-262
(1995)"
277..1800
/note="transmembrane protein"
/codon_start=1
/product="L amino acid transporter-1 LAT-1"
/protein_id="AAB93541.1"
/db_xref="GI:2731773"
/translation="MAADSVKRRSGASKTEEDRQAEKMLHONGNAEPKSGDAAY
ELQRTITVSKSGDYAVVLEVGALPAFLKLWVLLIIRPSOYVALVATYLLKPV
PPTCPVDDAAKIVACILILLTALNCYSVKAATVQDREAAKLLALLIILGFVQ
LGKGVDELKPERFEGSTNVGQWVALYSLGFAYGWNLYNFVVEEMIEPKNLR
AIIISMPIVTLVYLTNLAYFTLTPEQMLNSEAVAVDFGNHGLVWAWIIPVFGLS
CFGSYNGSLFTSSRLFFVGAREGHLPSLLAMHPRLITPMPSLIFTCAMTLLYAFSD
IPSVINFSFFNWLCAIIGMMMLRYKKPELRPIKVNILLPIFFILACILIVVS
FYMTPECGIGFIILLTGVPVYFFGVWQNKPDWLLHSHSSTALLQKMEAVPOES"
BASE COUNT        546 a 561 c 544 g 612 t
ORIGIN
Query Match      16.4%; Score 367; DB 5; Length 2263;
Best Local Similarity 55.4%; Pred. No. 6.7e-76;
Matches 734; Conservative 0; Mismatches 585; Indels 6; Gaps 1;
QY 340 GTCCAGCTGAAGAGGAATCTACTTTACGTAGGGGAGTCTCCATTATCATTTGGCACCATC 399
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 406 GTAGAGCTGCAAGCCACCATCTCTGTCATATGGGTGGCTATTATAGTGGGCCACCATC 465
QY 400 ATTTGAGCAGGAATCTTCTCTCTAAGGCGTGCTCCAGACACAGGGCAGCGTGGGC 459
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 466 ATCGGCTCGGGAATCTTCGTACACCCACCCCGGTGTGCTACGAGAGCGGGCTCCCGGG 525
QY 460 ATGCTCTGACCATCTGGAGGTGTGGGGTCTGTCTACTATTTGGAGCTTTGTCTTAT 519
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 526 CTGTGCTTTTGTATGGCGGTGTGCGGCCCTCTCTCCATTGTGGTCTGTGTTTAA 585
QY 520 GCTGAATGGGAACACTATAAGAAATCTGGAGGTCAATACACATATATTTTGAAGTC 579
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 586 GCCGAATGGGTACCAACTCTCCAAGTCTGGGGGGGACTATGCCTATGTGCTGGAGGTC 645
580 TTTGGTCCATTACAGAGCTTTTGTACGAGTCTGGTGGAACTCTCTCAATAATACCCCTGCA 639
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
646 TACGGTGGCTGCCCGCTTCTCTAAAGCTCTGGTGGAGCTGTCTCATCCGGCTCG 705
QY 640 GCTACTGCTGTATATCCCTGGGATTTGGAGGTACATCTCTGGAACCATTTTATTCAA 699
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 706 TCGCAATACATACGTGGTCTGGTGTCTGCTACCTACCTGCTCAAGCCGCTCTTCCCTACC 765
QY 700 TGTGAATCCTTGAACCTTTGGATCAAGCTCATTACAGCTGTGGGCATCACTGTAGTGATG 759
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 766 TGCCCCGTGCCCGATGATGCGGCAAGATCTGGCATGCCCTTTGTATCTTGTCTGCTGACT 825
QY 760 GTCTTAATAGCATGAGTGTGCTGAGCGCGCCGATCCAGATTTTCTTAACTTTTTCG 819
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 826 GCGATAAATGCTATAGTGTGAAAGAGCTACCAAGGTACAGATCGATTCGCGCAGCT 885
QY 820 AAGCTACAGCAATCTGATAATTAATAGTCCCTGGAGTTATGACGTAATTAAGGTCAA 879
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 886 AAATCTTGGCCCTTTTACTGATCATCATTTCTTGGTGTTCACCTGGGGAAGGGGC 945
QY 880 ACCGAGAACTTTAAAGAGC-----CCTTTTCAGGAAGAGATTCAGATATTACGCGGTG 933
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 946 GTGGAGGATCTGAGCCAGAAACGTTCTTTGAAGAACTAGCACCAATGTGTGCTAGTG 1005
QY 934 CCATGCTCTTTTATTTATGGAATGATGATGCTGGTGGTGTTCACCTCAACTTTTGT 993
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1006 GTCTGCAATGTACAGTGGCTTTTGGCCTATGAGGATGGAATCTTGAACCTTTGT 1065
QY 994 ACTGAAGAAGTAGAAACCTTGAAACCAATTTCCCTTTCGCAATATGATATCATATGCGC 1053
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1066 GTTGAAGAGATGATTGAACCTTACAGAATCTGCTCGAGCCATCATTTCCATGCC 1125
QY 1054 ATTTGACCATTTGCTATGCTGACAAATGTGSCCTACTTTACGACCAATTAATGCTGAG 1113
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1126 ATTGTAACCTCGGTATGTTTACTTAATCTGTCATATTTTCACTACTCTGACTCCAGAA 1185
QY 1114 GAGCTGCTGCTTCAAATGCAAGTGGCAGTGGCCTTTTCTGAGCGGCTACTGGAAATTC 1173
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1186 CAGATGCTGAATCAGAAGCTGTGGCTGTTCACCTTTGAAATTAACCATCTGGGGTCATG 1245
QY 1174 TCATTAGAGTTCGGATCTTTGTTGGCTCTCTGCTGTTTGGTCCATGAACGGTGGTG 1233
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1246 GCCTGGATTTATACAGATTTTGTGGATTTGTCATGCTTTGTTTTCAGTAAACGGCTCC 1305
QY 1234 TTTCTGCTCCAGGTTATCTATGTTGGTCTGAGAGGGTCACTTCCAGAAATCCTC 1293
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1306 TTTACATCTCCAGACTTTCTCGTGGAGCCAGAGAAGGGCATCTTCTCTTCTTACTT 1365
QY 1294 TCCATGATTCATGTCGCAAGACACACTCTCTACAGCTGTTATTTGTCACCCCTTG 1353
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1366 GCAATGATCCACCCCAAGGCTTCTTACTCCAATGCCGTCACTATATTCACCTTGGCCATG 1425
QY 1354 ACAATGATAATGCTTCTCTGAGAGACTCGACAGTCTTTTGAATTTCTTCAGTTTGGC 1413
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1426 ACGTCTCTATGCTTTCTCTGATGACATCTCTCTCTAATCAACTTCTTTCAGCTCTTT 1485
QY 1414 AGGTGGCTTTTATTTGGCTGGCAGTTGCTGGGCTGATTTATCTTCATACAAATGCCCA 1473
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1486 AACTGGCTTTGCGTAGCCCTGGCTATCATTTGTTATGTTGGCTTCGTTACAGAARACCA 1545
QY 1474 GATATGATCTCTCTTTCAAGGTGCCACTGTTCATCCAGCTTTTGTTCCTTCACATGC 1533
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1546 GAACCTAGAGAGGCTATCAAGGTTTAAATTTTGTGCTTATTTTTCATCTCTGGCCTGC 1605
QY 1534 CTCCTTATGTTGCCCTTTCCCTCTATTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1593
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1606 ATCTTCTCATCTGATGTCTATTTCTACATGACACAGTGTGAGTGGAAATGGATTTATC 1665
QY 1594 ATCACTCTGACTGGAGTCCCTCGCTATTATCTCTTTATATATGGGACAAAGAACCCAGG 1653
|| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 1666 ATTATTTGACTGGTGTCTCTACTCTTTTGGAGTCTGGTGGCAGATAAACCCGAC 1725
QY 1654 TGGTT 1658

```


Masiello, C., Maskeri, B., Mastrian, S. D., McCloskey, J. C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P. J., Touchman, J. W., Tsurgeon, C., Vogt, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E. D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 62 Row: h Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507052.

FEATURES

source

Location/Qualifiers
1..3448

/organism="Homo sapiens"
/db_xref="LocusID:9057"
/db_xref="taxon:9606"
/clone="MGC:40002 IMAGE:5214340"
/tissue_type="Blood, adult leukocytes"
/clone_lib="NIH_MGC_118"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
183..1730

CDS

/codon_start=1
/product="solute carrier family 7 (cationic amino acid transporter, y+ system), member 6"
/series_id="AAH28216.1"
/db_xref="GI:20381372"
/translation="MEAREPRTPTVHLVPNTSOSVEEDYSSPPORSETMQLKKE
ISLLNGSLVVGNIIGSGIFVSPKGVLVHTASVXGSLIIVWAIGGLFVVGALCYAELG
TTITKSGASVYILEAFGFIRLWVSLVVEPTGQALIAITFANYLIQSPFSCD
PPYLACRLAAACICILTFVNCAYRWKGTQVDTFYAKVALIAIIVMLVQLQGH
SEHFOADFGSSWMDNLSALYSALFSYSGWDTLNFVTEENRPNRPLAIGISMP
IVTLIYILTNVAYITLSDVLSDAVATPADOTFGMESWTIPTAVALSCFGLNA
SIPASRLPVGSRGHLPLDLSMIHIERPTLPALLFNCTMALIYILVEDVQLINY
FSPSYWFFVLSVGLYLRWKEPRRLKLSVFPPIVFCISVFLVLPVLPDTIN
SLGIGIALGSLVFPFYGVLPYLRNPLFIRNVLAAITRTQQLCPVLTELDVAEEK
KDRKTD"

BASE COUNT 782 a 861 c 846 g 959 t

ORIGIN

Query Match 16.9%; Score 379.2; DB 9; Length 3448;
Best Local Similarity 55.9%; Pred. No. 8.7e-79;

Matches 720; Conservative 0; Mismatches 568; Indels 0; Gaps 0;

Qy 334 GAGAAAGTCAGCTGAAGAGAAAGTCACTTACTGAGGGAGTCTCCATTATCATTTGGC 393
Dy 291 GAAATATGAGCTGAAGAGAGATCTCCCTGCTGATGGGTGACCCCTGTTGGTGGC 350
Qy 394 ACCATCATGAGCAGGAATCTTCATCTCCTAAAGGCGTGTCTCCAGAACACGGCAGC 453
Dy 351 AACATGATCGGCTCAGGGATCTTTGTCTCACCACCAAGGCTGTGCTGTGACACTGCTCC 410
Qy 454 GTGGCATGCTCTGACCATCTGACAGTGTGGGGTCTCTGTCACATTTGGAGCTTTC 513
Dy 411 TATGGGATGTCAGTGTGTGGGCACTTGTGGGCTCTCTCTGTTGGTGGGCTT 470
Qy 514 TCTTATGCTGAATGGGAACACTATAAAGAAATCTGAGGTCAATTACACATATATTTC 573
Dy 471 TGTATGAGAGCTGGGACCAACATCACCAGTCGGGACCAAGCTACGCTTATATCTA 530
Qy 574 GAAGCTTTTGGTCCATACCAGCTTTTGTACGAGTCTGGGTGGAACCTCTCATATAACGC 633
Dy 531 GAGGCTTTGGGGCTTCATTGCTTCATCCGCTGTGGGTCTCACTGTAGTTGTGAG 590
Qy 634 CCTGCACCTACTGCTGTGATATCCCTGGGATTTGGAGCTACATCTCGGAACATTTT 693
Dy 591 CCCACCGCTCAGGCCCATCATCGGCATCACCTTTGCCAACTACATCATCCAGCGCTCTC 650
Qy 694 ATTCATGTTGAATCCCTGAACTTGGCATCAAGCTCATACAGCTGTGGGCATACCTGTA 753
Dy 651 CCCAGCTGTATCCCATACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 710
Qy 754 GTGATGCTCTCAATAGCATGAGTGTGAGCTGAGGCGCCGCGATCCAGATTTCTTAAAC 813

Db 711 CTGACATTTGTGAACCTGTGCCTATGTCAAGTGGGGCACACGCTGTGCAGGACACGTTCACT 770
Qy 814 TTTTGAAGCTCACACCAATTTCTGATTAATATAGTCCCTGGAGTTATGACGCTAATATA 873
Dy 771 TAGCCCAAGGTCGTAGCGCTCAATGCCATCATGTGCAAGGGCTTTGTTAAACTGTGCCAG 830
Qy 874 GGTCAAAACGACGAACCTTTAAAGACGCTTTTTCAGGAAGAGATTTCAAGTATTAACGGGTG 933
Dy 831 GGACACTCTGAGCATTTCAGACGCTTTGAGGGTCTCTCTGGGACATGGGAAACCTC 890
Qy 934 CCACCTGGCTTTTATATGGAATGTATGCAATATGCTGGCTGTTTACCTCAACTTTGTT 993
Dy 891 TCTCTGCCCTCTACTCTGCCCTCTCTCTTACTCAGGTGGGACACCCCTTAATTTTGA 950
Qy 994 ACTGAAGAAGTAGAAGAACCTTGAAGAACCAATCCCTTGCATATGATATATCCATGGCC 1053
Dy 951 ACAGAAGAAATCAAAACCCAGAGAAATTTGCCCTTGGCCATTTGGGATTTCTATGGCA 1010
Qy 1054 ATTGTCAACCATTTGGCTATGTCTGACAAATGTGGCTACTTTTACGACCAATTAATGCTGAG 1113
Dy 1011 ATTTGACGCTCATCTACATCTGACCAATGTGGCTATTACACAGTGTGACATTTCA 1070
Qy 1114 GAGCTGCTCTTCAAAATCAGTGGCAGHAGCTTTTCTGAGCGCTACTTGGGAAATTC 1173
Dy 1071 GATGCTCTTAGCAGTATGCTGTGGCTGTGACATTTGCTGACACAGCTTTTGGCATGTT 1130
Qy 1174 TCATTAGCACTTCCGATCTTTTGGCCCTCTCTGCTTGGCTTCCATGAACGGTGGTGTG 1233
Dy 1131 AGCTGGACCATCCCATTTGCTTGGCTGTCTGCTTGGGGGCTCAATGCATCCATC 1190
Qy 1234 TTTGCTGTCTCCAGGTATTCTATGTGTGCTCGAGAGGGTCACCTTCCAGAAATCTCTC 1293
Dy 1191 TTTGCTTCAAGGTGTGTTCTGTTGGGCTCCCGGAGGGCCACTACCGGACCTCTCTG 1250
Qy 1294 TCCATGATTCATGTCGCGAAGCACACTCTCTACGAGCTGTATTGTTTGGCACCTTTG 1353
Dy 1251 TCCATGATCCCATTCAGCGTTTACACCTATCCCTGCTTACTGTTCAATTTGCACCATG 1310
Qy 1354 ACAATGATTAATGCTCTCTCGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGC 1413
Dy 1311 GCACATCATCTACTCATCTGTTGGAGGATGTTTTCAGCTTATCAACTACTTCACTGCTCAGC 1370
Qy 1414 AGGTGCTTTTATTTGTTGGCTGCGAGTGTCTGGCTGATTATCTTCGATACAAATGCCCA 1473
Dy 1371 TACTGTTCTTCTGTTGGCTGCTGTGTTGGACAGCTCTACTCCGCTGGAGAGGCC 1430
Qy 1474 GATATGATGCTGCTTTCAAGGTGCGACTGTTTCATCCAGCTTTTGTTCCTTCATATGC 1533
Dy 1431 AAGCGGCCCGGCTCTCAAGCTGAGCGTGTTTTCCCATCGTCTGTTCTGCAATATGCTCC 1490
Qy 1534 CTCTTCATGTTGCCCTTTCCCTCTATTTCGGACCCATTTAGTACAGGATTTGGCTTCGTC 1593
Dy 1491 GTGTTCTGTTGATAGTGGCCCTCTCTCACTGACACCACTTAATTCCTCATTTGGCATCGG 1550
Qy 1594 ATCACTCTGACGTGAGTCCCTGCGTATT 1621
Dy 1551 ATTGCCCTTCTGGAGTCCCTTCTACT 1578

RESULT 12

D87432

LOCUS D87432 Human mRNA for KIAA0245 gene, complete cds.
DEFINITION D87432
ACCESSION D87432.1 GI:1665758
VERSION KIAA0245.
KEYWORDS

SOURCE Homo sapiens male bone marrow Myeloblast cell line:KG-1 CDNA to

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Nagase, T., Seki, N., Ishikawa, K., Ohira, M., Kawarabayashi, Y.,

misc_feature 4137..4527 /note="match to EST AI267688 (NID:g3886855) aq93c10.x1"
misc_feature 4153..4433 /note="match to EST C00538 (NID:g1432768)"
misc_feature 4233 /note="match to EST BI497113 (NID:g15336457)"
misc_feature 4359..4746 /note="match to EST AA227760 (NID:g1849333) zr56a07.r1"
misc_feature 4404..5178 /note="match to EST BG291471 (NID:g13049386)"
misc_feature 4412..4619 /note="similar to Homo sapiens EST BF154977 (NID:g11050160)"
misc_feature 4657..5028 /note="match to EST R20307 (NID:g774941) y935a09.r1"
misc_feature 4727..5195 /note="similar to Bos taurus EST AW669076 (NID:g7525590)"
misc_feature 4740..5803 /note="similar to EST BM547723 (NID:g18781736)"
misc_feature 4859..5265 /note="match to EST AW901105 (NID:g8065310)"
misc_feature 5002..5687 /note="match to EST BG254117 (NID:g12763933)"
misc_feature 5004..5213 /note="match to EST BF084219 (NID:g10878049)"
misc_feature 5074..5447 /note="match to EST N80420 (NID:g1243121) yz92e06.r1"
misc_feature 5255..6002 /note="match to EST BI869486 (NID:g15043159)"
misc_feature 5442..6151 /note="match to EST AW961203 (NID:g8150887)"
misc_feature 5497..5830 /note="match to EST AI540897 (NID:g4458270)"
misc_feature 5782..6124 /note="match to EST AA309186 (NID:g1961511)"
misc_feature 5875..6701 /note="match to EST BI334972 (NID:g15019629)"
misc_feature 5995..6331 /note="match to EST AA471165 (NID:g2199022)"
misc_feature 6000..6124 /note="match to EST AW996896 (NID:g8257130)"
misc_feature 6032..6836 /note="similar to EST BM474943 (NID:g18523985)"
misc_feature 6140..6518 /note="match to EST AW996897 (NID:g8257131)"
misc_feature 6444..6885 /note="match to EST AA375957 (NID:g2028275)"
repeat_region 6878..7187 /rpt_family="Alu"
misc_feature 6899..7362 /note="match to EST AA488687 (NID:g2216118) ab38f03.s1"
misc_feature 6970..7746 /note="match to EST BE739933 (NID:g10153837)"
misc_feature 7099..7815 /note="match to EST BF381953 (NID:g11363256)"
misc_feature 7124..7125 /note="match to EST AW996897 (NID:g8257131)"
misc_feature 7174..7495 /note="match to EST BF927528 (NID:g12325656)"
misc_feature 7253..7884 /note="similar to Homo sapiens EST BG499179 (NID:g13460696)"
misc_feature 7568..7949 /note="match to EST AA307001 (NID:g1959330)"
misc_feature 7613..8332 /note="similar to EST BM706727 (NID:g19019985)"
repeat_region 7836..8011 /rpt_family="Achoho"
misc_feature 8035..8572

Query Match 23.6%; Score 529.2; DB 9; Length 64196;
Best Local Similarity 99.4%; Pred. No. 2.9e-114;
Matches 531; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1671 AGAGAAAATAACAGAACATTTACAAATAATACTGGAAGTTGTACCAGAGAGATAAGTT 1730
Db 11755 AGAGAAAATAACAGAACATTTACAAATAATACTGGAAGTTGTACCAGAGAGATAAGTT 11696
QY 1731 ATGAATAATGAGACTTGGAGACTTGGCAATCTGCCCAAGGGAGAGACACAAAATAGGGATT 1790
Db 11695 ATGAATAATGAGACTTGGAGACTTGGCAATCTGCCCAAGGGAGAGACACAAAATAGGGATT 11636
QY 1791 TTTACTTCATTTTCTGAAAGTCTAGAGAAATTACAACCTTTGGTGATATAAACAAGAGAGTCA 1850
Db 11635 TTTACTTCATTTTCTGAAAGTCTAGAGAAATTACAACCTTTGGTGATATAAACAAGAGAGTCA 11576
QY 1851 GTTATTTTATTCATATATTTTAGCATATTTTGAACCTTAATTTCTAAGAAATTTAGTTATATA 1910
Db 11575 GTTATTTTATTCATATATTTTAGCATATTTTGAACCTTAATTTCTAAGAAATTTAGTTATATA 11516
QY 1911 CTCTATGATGATATAGAAAGTGAATATGCAATTTATTTCTATGAGTCGACAAATTTCTTGAGT 1970
Db 11515 CTCTATGATGATATAGAAAGTGAATATGCAATTTATTTCTATGAGTCGACAAATTTCTTGAGT 11456
QY 1971 CTCTGATACCTACCTATTGGGGTTAGGAGAAAAGACTAGACAATTTACTATGTGGTCAATTC 2030
Db 11455 CTCTGATACCTACCTATTGGGGTTAGGAGAAAAGACTAGACAATTTACTATGTGGTCAATTC 11396
QY 2031 TCTACAAATATGTTAGCACGCGCAAGAACTTCAAAATTTGAAGACTGAGATTTTCTGTGA 2090
Db 11395 TCTACAAATATGTTAGCACGCGCAAGAACTTCAAAATTTGAAGACTGAGATTTTCTGTGA 11336
QY 2091 TATATGGGTTTTGTAAGATGGTTTTACACACTACAGATGCTCTATCTGTGAAAAGTGTT 2150
Db 11335 TATATGGGTTTTGTAAGATGGTTTTACACACTATAGATGCTCTATCTGTGAAAAGTGTT 11276
QY 2151 TTCAATTTCTGAAAAAAGCATACATCATGATTATGGCAAAAGAGAGAGAGAGTA 2204
Db 11275 TTCAATTTCTGAAAAAAGCATACATCATGATTATGGCAAAAGAGAGAGAGAGAGTA 11222

RESULT 11
LOCUS BC028216
DEFINITION Homo sapiens, solute carrier family 7 (cationic amino acid transporter, y+ system), member 6, clone MGC:40002 IMAGE:5214340, mRNA, complete cds.
ACCESSION BC028216
VERSION BC028216.1 GI:20381371
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3448)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NTH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland.
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghig, P., Hansen, N., Ho, S.-L., Karlins, E., Laric, P., Legaspi, R., Maduro, O.L.,

```

QY 1974 TGATACCTACTATTGGGTTAGGAGAAAGACTAGACAACTACTAT-----GTGGTC 2026
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 2057 GTGCTATCTGCTGGGGGAAAGAAAAAACCTAAGGCTTTGCTACTGGGCTATCA 2116
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
QY 2027 ATTCTTACACATATGTTAGCAGCGCAAGAACCTTCAAAATTGAAGACTGAGATTTTC 2086
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 2117 TTCTCTACACACGCTTTCTTATTATGACTGAGAACCCTTGAACAGACCAACAAATGGTTTC 2176
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
QY 2087 TGTATATATGGGTTTGTAAAGATGCTTTTACACACT 2123
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 2177 TGTATATATGAGGCTCTATAACATAGCTTTACCTACT 2213
      |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

RESULT 10
AC110804/c
LOCUS      AC110804      64196 bp      DNA      linear      PRI 30-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-725C19 from 4, complete sequence.
ACCESSION  AC110804
VERSION    HTG.
KEYWORDS   AC110804.3  GI:19570222
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 64196)
AUTHORS   Sulston, J.E. and Waterston, R.
TITLE     Toward a complete human genome sequence
JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE   99063792
PUBMED    9847074
REFERENCE  2 (bases 1 to 64196)
AUTHORS   Pearman, C., Kozlowski, A. and Dignan, G.
TITLE     The sequence of Homo sapiens BAC clone RP11-725C19
JOURNAL   Unpublished (2001)
REFERENCE  3 (bases 1 to 64196)
AUTHORS   Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (15-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  4 (bases 1 to 64196)
AUTHORS   Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (06-MAR-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  5 (bases 1 to 64196)
AUTHORS   Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (21-MAR-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  6 (bases 1 to 64196)
AUTHORS   Waterston, R.
TITLE     Direct Submission
JOURNAL   Submitted (30-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Mar 21, 2002 this sequence version replaced gi:19073866.
            -----
            Center: Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu
            -----
            Summary Statistics
            -----
            Center project name: H_NH0725C19
            -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-8P17, 2000 bp overlap. Actual start of this clone is at base position 185268 of RP11-393A12; actual end is at base position 64196 of RP11-725C19.

Polymorphisms exist between AC013365 and AC110804.

FEATURES	source
Location/Qualifiers	1. .64196
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone="RP11-725C19"
	/clone_lib="RPCI-11"
repeat_region	1. .39
	/rpt_family="AT-rich"
repeat_region	162. .182
	/rpt_family="AT-rich"
repeat_region	312. .607
	/rpt_family="Alu"
misc_feature	369. .505
	/note="match to EST BC257959 (NID:gl2767775)"
repeat_region	877. .909
	/rpt_family="AT-rich"
repeat_region	1113. .1536
	/rpt_family="L2"
repeat_region	2400. .2427
	/rpt_family="AT-rich"
repeat_region	2941. .2984
	/rpt_family="L2"
repeat_region	3413. .3435
	/rpt_family="(TA)n"
misc_feature	3829. .4216
	/note="match to EST H06135 (NID:g869687) y177b12.s1"
misc_feature	3829. .4170
	/note="match to EST BI497113 (NID:gl5336457)"
misc_feature	3831. .4157
	/note="match to EST AA639698 (NID:g2563477) np60f09.s1"
misc_feature	3835. .4360
	/note="match to EST AW237145 (NID:g6569534) xm66b03.x1"
misc_feature	3836. .4376
	/note="similar to EST BM791320 (NID:gl9139552)"
misc_feature	3837. .4405
	/note="similar to Bos taurus EST BF539797 (NID:gl5380907)"
misc_feature	3844. .4610
	/note="match to EST BI857800 (NID:gl5998547)"
misc_feature	3933. .4394
	/note="match to EST AI694383 (NID:g4971723) wd83b04.x1"
misc_feature	4105. .4843
	/note="match to EST BG177619 (NID:gl2684322)"

/gene="xct"
/function="anionic amino acid transport"
/codon_start=1
/evidence=experimental
/product="cystine/glutamate transporter"
/protein_id="BAA77220.1"
/db_xref="GI:4689081"
/translation="MVRKPVVATISKGYLQGNMGRPLSMGDEPPGQKVVLLKKK
TLKRSGLIIGTVLSEGFIFSPKILQNTGSMGLVFWMSACVLSLFGALSVAELGT
SIRKSGHYTVLEVFGLPILAFVRVVELLIVRGATAVISLAFGRVILPFPFIOCEI
PELAIKLVAVGIIIVVNLSTSVWSARIQIELTCKLPAILLIIVPGVILIKGOT
HFKDAFSGRDTSLMGLPLAFIYGYAYAGWFYINFTVEEDNPEKTIPLAICISNAI
ITGVYVNTNAYFTISAEELLQSSAVAVTFSERLKGLFSLAVPIFVALSCFSGMNGG
VFAYSRIFYVASRGLHPEILSMHVHKHPLPAVIVLHPLTMVLFSGDLYSLNPL
SFARWLEMLAVAGLILVLRKPDHPRPKVPLFIPALFSTCLEMVLVLSYSDPFT
GVGELITLGVPAIYLFIVWDKKPKWFRRLSDRITRLQIILEVVPEDSKEL"
2216

polyA_site

/gene="xct"

/note="34 a nucleotides"

BASE COUNT 542 a 526 c 509 g 639 t

ORIGIN

Query Match 54.0%; Score 1208.6; DB 10; Length 2216;
Best Local Similarity 79.1%; Pred. No. 1.3e-274;
Matches 1548; Conservative 0; Mismatches 324; Indels 85; Gaps 6;
QY 182 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGTTACTGCGAGGAAATGTTAACGGGA 241
DB 327 GAAAGCAGTTGTGGCCACCATCTCCAAAGAGGTTACTGCGAGGCAATATGACGGGA 386
QY 242 GGCTGCTTCCCTGGGCAACAGAGGCCACCTGGGCGACGCCCTTTTCAGGAAGAGACG 301
DB 387 GGCTGCCCTCCATGGGGACCAAGAGCCACCTGGGCAG----- 424
QY 302 CCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAGTGCAGCTGAAGAGGAAGTCA 361
DB 425 -----GAGAAGGTAGTCTCGAAAAAGAGATCA 452
QY 362 CTTTACTGAGGGAGTCTCATATATCATTTGGCACCACCATATTTGAGCAGGAATCTTCATCT 421
DB 453 CTTTGTGAGGGGGTCTCCATCATCATCGGCACCGTCTATCGATCAGGCATCTTCATCT 512
QY 422 CTCTTAAGGGCGTGTCCAGACACGGGCGAGGTGGGCAATGCTCTGACCATTCTGGACGG 481
DB 513 CCCCCAAGGCGATACCTCCAGAACACGGGCGAGGTGGGCATGTCCCTGGTCTTCTGGCTGT 572
QY 482 TGTGTGGGTCTGTCTCAGTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAA 541
DB 573 CTTGTGGAGTACTGTGCACCTTTTGGAGCCCTGTCTTATGCAAAATAGTACAGCATAA 632
QY 542 AGAAATCTGGAGTCTATTACATATATTTTGGAGTCTTTGGTCCATTACCAGCTTTTG 601
DB 633 AGAAATCTGGTGTCTATTACATATCTTGGAGTCTTTGGTCTTTTGGTCTTTGCTTTG 692
QY 602 TAGAGTCTGGGTGGAACTCTCATATACGCCCTGCACCTACTGCTGTGATATCCCTGG 661
DB 693 TTAGAGTCTGGGTGGAACTCTCGTATACGCCCTGGAGTCTTTGGTCTTTTGGTCTTTG 752
QY 662 CATTTGGAGCTACATCTTGGAAACCATTTTATTCAAATGTGAATTCCTGAACTTGGCA 721
DB 753 CATTTGGAGCTACATCTTGGAAACCATTTTATTCAAATGTGAATTCCTGAACTTGGCA 812
QY 722 TCAAGCTCATTTACAGCTGTGGGCAATCTGTAGTGTGTCCTTAAATAGCATGAGTGCA 781
DB 813 TCAAGCTCGTACAGCTGTGGGCAATCTGTGTGTGATGTCTTAAATAGCATGAGTGCA 872
QY 782 GCTGGAGCGCGGATCCAGATTTTCTTAACCTTTTGGAGCTTCACAGCAATTCATGATA 841
DB 873 GCTGGAGTGGCGGATCCAGATTTTCTTAACCTTTTGGAGCTTCACAGCAATTCATGATA 932
QY 842 TTATAGTCCCTGGAGTTATGCACTAATTAAGGCTCAACGCAAGAGCTTTTAAAGAGCGCT 901
DB 933 TTATAGTCCCTGGAGTTATGCACTAATTAAGGCTCAACGCAAGAGCTTTTAAAGATGCAT 992

QY 902 TTTTCAGGAAGAGATTCAAAGTATTTACGGGTTGCCACTGGCTTTTATTTATGGAATGATG 961
DB 993 TTTTCAGGAAGAGAGACAAAGTCTAATGGGTTGCCCTTTGCTTTTATTTATGGAATGATG 1052
QY 962 CATATGCTGCTGGTGTTCACCTCAACTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAA 1021
DB 1053 CATATGCTGCTGGTGTTCACCTCAACTTGTACTGAAGAAGTAGAAAAACCCCTGAAAAA 1112
QY 1022 CCATTCCTCCCTTGCATATGTATATCCATGCCCATTTGTCAACATTGGCTATGTCTGACAA 1081
DB 1113 CCATTCCTCCCTTGCATATGTATATCCATGCCCATTTGTCAACATTGGCTATGTCTGACAA 1172
QY 1082 ATGTGGCCTACTTTACAGCATTAAATGCTGAGAGCTGCTGCTTTCAATATGCAAGTGGCAG 1141
DB 1173 ACGTGGCCTATTTTACCACCATCAGTGGGAGGAGCTGTGTCAGTCCAGCGCGTGGCGG 1232
QY 1142 TGACCTTTTCTGAGCGGCTACTGGGAATTTCTCATTAGCAGTTCGGATCTTTTGTGGCCC 1201
DB 1233 TGACCTTTTCTGAGCGGCTGCTGGGAATTTCTCATTAGCAGTTCGGATCTTTTGTGGCCC 1292
QY 1202 TCTCTGCTTTGGCTCCATGAACGGTGTGTGTGTCTGCTCTCCAGGTTATTTCTATGTTG 1261
DB 1293 TCTCTGCTTTGGCTCCATGAACGGTGTGTGTGTCTGCTCTCCAGGTTATTTCTACGTCG 1352
QY 1262 GGTCTCAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGTCGCGCAAGCACACTC 1321
DB 1353 CATCTCAGAGGGGACCTTCCGGAATCCTCTCTATGATTCATGTCACCAAGCACACTC 1412
QY 1322 CTCTACAGCTGTTTATTTGTTTGCACCTTTTGACAATGATAATGCTCTTCTCGGAGACC 1381
DB 1413 CTCTGCGAGCTGTTTATTTGTTTGTGATCCTCTGACATGTTGTTGATGCTCTTCTCGGAGACC 1472
QY 1382 TCGACAGCTTTTGAATTTCTCAGTTTTCAGAGTGGCTTTTATTTGGCTGGCAGTTG 1441
DB 1473 TCTATAGTCTTAAATTTCTCAGTTTTCAGAGTGGCTTTTATTTGGCTGGCAGTTG 1532
QY 1442 CTGGGCTGATTTATCTCGATACAAATGCCAGATATGCATCGCTCTTTCAGAGTGGCAC 1501
DB 1533 CAGGACTGATTTATCTTCGATACAAAGCCAGATATGCATCGCTCTTTCAGAGTGGCTC 1592
QY 1502 TGTTCATCCAGCTTGTGTTTCTTCATGCTCTTCATGCTTTCATGCTTTCCTCTCTATT 1561
DB 1593 TCTTCATCCCGGCACTATTTTCTTCCCTGCTTCCATGCTTTCATGCTTTCCTCTCTTACT 1652
QY 1562 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCTGACCTGAGTGGAGTCCCTGCGTATT 1621
DB 1653 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCTGACCTGAGTGGAGTCCCTGCGTATT 1712
QY 1622 ATCTCTTTTATTTATGGGACAAAGAACCCAGTGGTTTAGAATATGTCAGAGAAATAA 1681
DB 1713 ATCTCTTTCATTTGATGGGACAAAGAACCCAGTGGTTTAGAATATGTCAGAGAAATAA 1772
QY 1682 CCAGAACATTTACAATTAATCTGGAAGTTGTACCAAGAAAGATA---AGTTATGAACATA 1738
DB 1773 CCAGAACATTTACAATTAATCTGGAAGTTGTACCAAGAAAGATA---AGTTATGAACATA 1832
QY 1739 ATGACTTTGAGATCTTGGCAATCTGCCCAAGGGAGACACAAAATAGGATTTTACTTTC 1798
DB 1833 AATGCATCAAAAGCTTGGCCATCTGCC-AGGATTTAGATACAAAATGGAATTTT----- 1886
QY 1799 ATTTTCTCAAGTCTAGAGAATTTACAATTTGGTGCATAAACAAAGAGTGCAGTTATTTT 1858
DB 1887 -----ATTCACAAAGAACACACAGCTTGGATGGATGAAGTAATCAGTTATCTC 1936
QY 1859 TATTTCATATTTTACCATATTTGCAACTTAATTTCTTAAGAAATTTAGTTTATAACTCTATGT 1918
DB 1937 TATTTCATATCTTACGCTATTTCAAATTAATTTTGTGACCACTTACTGTTAACTCCATGT 1956
QY 1919 AGTTATAGAAAGTGAATATGCAAGTT-----ATTCATGAGTCGCGCAATTTCTGAGTCTC 1973
DB 1997 ATTTGTAGCAAGCTAATATGCAAGTGCATACAGTGGAGGCAAGCTCACAGTTCTTCTGAGTCTA 2056

Query Match SFARWLEIGLAVAGLIYLRKCPDMHRPEKVPILPALFSFTCLPMVALSYSDPEST
 Best Local Similarity 96.0%; Pred. No. 1.4e-313;
 Matches 1441; Conservative 0; Mismatches 6; Indels 54; Gaps 1;
 BASE COUNT 505 a 418 c 435 g 642 t
 ORIGIN

182 GAAGAGCTGTGTCTCCACCATCTCCAAAGAGGTTTACCTGCAGGGAATGTTAAGCGGA 241
 143 GAAGAGCTGTGTCTCCACCATCTCCAAAGAGGTTTACCTGCAGGGAATGTTAAGCGGA 202
 242 GGCTGCCCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCGCTTTTCAGGAGAGAGCG 301
 203 GGCTGCCCTTCCCTGGGCAACAGGAGCCACCTGGGAG----- 240
 302 CTTTTTCAGGAGAGAGCGCTTTTCAGGAGAGAGAGAGTGCAGCTGAAGAGGAAAGTCA 361
 241 -----GAGAAAGTGCAGCTGAAGAGGAAAGTCA 268
 362 CTTTACTGAGGGAGTCTCCATTCATTCGGCACCATCATTCGAGCAGGAATCTTCATCT 421
 269 CTTTACTGAGGGAGTCTCCATTCATTCGGCACCATCATTCGAGCAGGAATCTTCATCT 328
 422 CTCTAAGGGCGTCTCCAGAACACAGCGGAGCGTGGGCATGTCTGTGACCATCTGGAGCG 481
 329 CTCTAAGGGCGTCTCCAGAACACAGCGGAGCGTGGGCATGTCTGTGACCATCTGGAGCG 388
 482 TGTGTGGGTCCTGTCTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGAAACAATATAA 541
 389 TGTGTGGGTCCTGTCTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGAAACAATATAA 448
 542 AGAAATCTGAGGTCATTCACATATATTTGGAAGTCTTTGGTCCATTTACAGCTTTTG 601
 449 AGAAATCTGAGGTCATTCACATATATTTGGAAGTCTTTGGTCCATTTACAGCTTTTG 508
 602 TAGAGTCTGGGTGGGAACTCCTCATATACGGCCCTCAGCTACTGCTGTGATATCCCTGG 661
 509 TAGAGTCTGGGTGGGAACTCCTCATATACGGCCCTCAGCTACTGCTGTGATATCCCTGG 568
 662 CATTGGAGCTACATCTGGAAACCAATTTTATTCAAATGTGAATFCCCTGAACCTGGGA 721
 569 CATTGGAGCTACATCTGGAAACCAATTTTATTCAAATGTGAATFCCCTGAACCTGGGA 628
 722 TCAAGCTCATTTACAGCTGTGGGCAATTAAGTGTAGTGTCTCTAAATAGCATGAGTGCA 781
 629 TCAAGCTCATTTACAGCTGTGGGCAATTAAGTGTAGTGTCTCTAAATAGCATGAGTGCA 688
 782 GCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCAGCAATTTCTGATAA 841
 689 GCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCAGCAATTTCTGATAA 748
 842 TTATAGTCTCTGGAGTTAGCATTAATTAAGTGTAAAGTCAACAGCAGCACTTTAAGAGCGCT 901
 749 TTATAGTCTCTGGAGTTAGCATTAATTAAGTGTAAAGTCAACAGCAGCACTTTAAGAGCGCT 808
 902 TTTAGGAAGAGATCAAGTATTACGGGTTGCCACTGGCTTTTATTTATGGAATGTATG 961
 809 TTTAGGAAGAGATCAAGTATTACGGGTTGCCACTGGCTTTTATTTATGGAATGTATG 868
 962 CATATGCTGGCTGGTTTACCTCACTTTGTTTACTGAAGAGTAGAAGCAACCTGAAAAA 1021
 869 CATATGCTGGCTGGTTTACCTCACTTTGTTTACTGAAGAGTAGAAGCAACCTGAAAAA 928
 1022 CCATTCCTCCCTGCAATATGATATCATGCGGATGTCACCATTTGGCTATGCTGACAA 1081
 929 CCATTCCTCCCTGCAATATGATATCATGCGGATGTCACCATTTGGCTATGCTGACAA 988
 1082 ATGTGCGCTTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTTCAAATGAGTGGCAG 1141
 989 ATGTGCGCTTACTTTACGACCAATTAATGCTGAGGAGCTGCTGCTTTCAAATGAGTGGCAG 1048

QY 1142 TGACCTTTTCTGAGCGGCTACTGGAAATTTCTCATTAGCAGTTCGGATCTTTGTTGCC 1201
 Db 1049 TGACCTTTTCTGAGCGGCTACTGGAAATTTCTCATTAGCAGTTCGGATCTTTGTTGCC 1108
 QY 1202 TCTCTGCTTTTGGCTTCCATGAACGCTGGTGTGTTTCTGCTCTCCAGGTTATTCTATGTTG 1261
 Db 1109 TCTCTGCTTTTGGCTTCCATGAACGCTGGTGTGTTTCTGCTCTCCAGGTTATTCTATGTTG 1168
 QY 1262 CTTCTCGAGAGGCTCACCTTCCAGAAATCTCTCCATGATTCATCTCCGCAAGCACACTC 1321
 Db 1169 CTTCTCGAGAGGCTCACCTTCCAGAAATCTCTCCATGATTCATCTCCGCAAGCACACTC 1228
 QY 1322 CTTCTACCAGCTGTTATTTGTTTTCACCCCTTTTGACAATGATTAATGCTCTCTCTGAGACC 1381
 Db 1229 CTTCTACCAGCTGTTATTTGTTTTCACCCCTTTTGACAATGATTAATGCTCTCTCTGAGACC 1288
 QY 1382 TCGACAGCTTTTGAATTTCTCAGTTTTCAGAGTGCTTTTATTTGGCTGGCAGTTG 1441
 Db 1289 TCGACAGCTTTTGAATTTCTCAGTTTTCAGAGTGCTTTTATTTGGCTGGCAGTTG 1348
 QY 1442 CTGGGCTGATTATCTTCGATACAAATGCCAGATATGATCATCTCTTCAAGGTGCCAC 1501
 Db 1349 CTGGGCTGATTATCTTCGATACAAATGCCAGATATGATCATCTCTTCAAGGTGCCAC 1408
 QY 1502 TGTTCATCCAGCTTTGTTTTCCTTCACATGCCTTTCATGCTGCTTCCCTCTCTATT 1561
 Db 1409 TGTTCATCCAGCTTTGTTTTCCTTCACATGCCTTTCATGCTGCTTCCCTCTCTATT 1468
 QY 1562 CGGACCCATTTAGTACAGGAGTTGGCTTCGTCATCCTCTGAGTCCCTCCGCTATT 1621
 Db 1469 CGGACCCATTTAGTACAGGAGTTGGCTTCGTCATCCTCTGAGTCCCTCCGCTATT 1528
 QY 1622 ATCTCTTTATTAATGAGCAAGAACCCAGGTGGTGTAGTAATATGTCAGAGAAAAATAA 1681
 Db 1529 ATCTCTTTATTAATGAGCAAGAACCCAGGTGGTGTAGTAATATGTCAGAGTTCTCTAG 1588
 QY 1682 C 1682
 Db 1589 C 1589

RESULT 9
 AB022345
 LOCUS
 DEFINITION Mus musculus xct mRNA for cystine/glutamate transporter, complete cds.
 ACCESSION AB022345
 VERSION J. 022345.1 GI:4689080
 KEYWORDS xct; cystine/glutamate transporter.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Sato, H., Tamba, M., Ishii, T. and Bannai, S.
 Cloning and expression of a plasma membrane cystine/glutamate exchange transporter composed of two distinct proteins
 J. Biol. Chem. 274 (17), 11455-11458 (1999)
 9223452
 2 (bases 1 to 2216)
 REFERENCE Sato, H. and Bannai, S.
 Direct Submission
 Submitted (13-JAN-1999) Hideyo Sato, University of Tsukuba,
 Institute of Basic Medical Sciences, Tennodai 1-1-1, Tsukuba,
 Ibaraki 305-8575, Japan (E-mail:hideyo-semd.tsukuba.ac.jp,
 Tel:81-298-53-3282, Fax:81-298-53-3039)
 FEATURES
 source 1. .2216
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 gene 1. .2216
 /gene="xct"
 cds 320. .1828

Db 429 CTCCTAAGGGCGTCTCCAGAACACGGGACGGTGGGCATGTCTCTGACCATCTCGACGG 488
QY 482 TGTGTGGGTCTGTACATATTGGAGCTTTGTCTTATGTCTGAATGGGAACAATATAA 541
Db 489 TGTGTGGGTCTGTACATATTGGAGCTTTGTCTTATGTCTGAATGGGAACAATATAA 548
QY 542 AGAAATCTGGAGTCTATACACATATATTTTGGAGTCTTTGGTCCATATACCAGCTTTTG 601
Db 549 AGAAATCTGGAGTCTATACATATATTTTGGAGTCTTTGGTCCATATACCAGCTTTTG 608
QY 602 TAGAGCTCTGGGTGGAACTCTCATATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 661
Db 609 TAGAGCTCTGGGTGGAACTCTCATATACGCCCTGCAGCTACTGCTGTGATATCCCTGG 668
QY 662 CATTTGGAGCTACATCTTGGAACTATTTTATTCATGTCAATGCTGAACCTGGTGGCA 721
Db 669 CATTTGGAGCTACATCTTGGAACTATTTTATTCATGTCAATGCTGAACCTGGTGGCA 728
QY 722 TCAAGCTCATTTACAGCTGTGGGCATTAACCTGTAGTGTGCTCTAAATAGCATGAGTGCA 781
Db 729 TCAAGCTCATTTACAGCTGTGGGCATTAACCTGTAGTGTGCTCTAAATAGCATGAGTGCA 788
QY 782 GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCAGCAATTCGTGATAA 841
Db 789 GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCAGCAATTCGTGATAA 848
QY 842 TTATAGTCCCTGAGTATGACGCTAATTAAGGTCACACGACGACCTTAAAGACGCT 901
Db 849 TTATAGTCCCTGAGTATGACGCTAATTAAGGTCACACGACGACCTTAAAGACGCT 908
QY 902 TTTTCAAGAGAGATTAAGTATTTACGCGTGTGCACTGGCTTTTATTTATGGAATGTATG 961
Db 909 TTTTCAAGAGAGATTAAGTATTTACGCGTGTGCACTGGCTTTTATTTATGGAATGTATG 968
QY 962 CATATGCTGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCTGAAAAA 1021
Db 969 CATATGCTGCTGGTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAAACCTGAAAAA 1028
QY 1022 CCATTCCTCTGCAATATGATATCATGCGCATTCGCACTGGCTGTGCTGCTGACAA 1081
Db 1029 CCATTCCTCTGCAATATGATATCATGCGCATTCGCACTGGCTGTGCTGCTGACAA 1088
QY 1082 ATGTGGCTACTTTACGACCAATTAATGCTGAGGAGTGTGCTTTTCAAAATGAGTGGCAG 1141
Db 1089 ATGTGGCTACTTTACGACCAATTAATGCTGAGGAGTGTGCTTTTCAAAATGAGTGGCAG 1148
QY 1142 TGACCTTTTCTGAGGCTACTGGAATTTCTCATATAGCAGTTCGATCTTTGTTGCC 1201
Db 1149 TGACCTTTTCTGAGGCTACTGGAATTTCTCATATAGCAGTTCGATCTTTGTTGCC 1208
QY 1202 TCTCTGCTTTGGTCTCCATGAACGGTGGTGTGCTGCTCCAGGTTATTTCTATGTTG 1261
Db 1209 TCTCTGCTTTGGTCTCCATGAACGGTGGTGTGCTGCTCCAGGTTATTTCTATGTTG 1268
QY 1262 CGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGCGCAAGCACACTC 1321
Db 1269 CGTCTCGAGAGGTCACCTTCCAGAAATCCTCTCCATGATTCATGCGCAAGCACACTC 1328
QY 1322 CTCTACAGCTGTTATTTGTTTTGCAACCTTTGCAATGATATGCTCTCTCGAGAC 1381
Db 1329 CTCTACAGCTGTTATTTGTTTTGCAACCTTTGCAATGATATGCTCTCTCGAGAC 1388
QY 1382 TCGACAGCTTTTGAATTTCTCAGTTTTCGAGGTTGGCTTTTATTTGGCTGGCAGTTG 1441
Db 1389 TCGACAGCTTTTGAATTTCTCAGTTTTCGAGGTTGGCTTTTATTTGGCTGGCAGTTG 1448
QY 1442 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCAATGCTCTCTTCAAGTCCAC 1501
Db 1449 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCAATGCTCTCTTCAAGTCCAC 1508
QY 1502 TGTTCATCCAGCTTTGTTTCTTCCATGACGCTCTTCATGCTGTTGGCTTTCCCTCTATT 1561
Db 1509 TGTTCATCCAGCTTTGTTTCTTCCATGACGCTCTTTCATGCTGTTGGCTTTCCCTCTATT 1568

QY 1562 CGGACCATTTAGTACAGGATTTGGCTTCGCTCATCTCTGACTGGAGTCCCTGCGTATT 1621
Db 1569 CGGACCATTTAGTACAGGATTTGGCTTCGCTCATCTCTGACTGGAGTCCCTGCGTATT 1628
QY 1622 ATCTCTTTATATATGGCAAGAAACCCAGGTGGTTTGTAGTAATATCTCAGAGAAATAA 1681
Db 1629 ATCTCTTTATATATGGCAAGAAACCCAGGTGGTTTGTAGTAATATCTCAGAGAAATAA 1688
QY 1682 CCAGAACATTTACAATATATCTGGAAGTTGTACCAGAGAGATAAGTTATGAACTAATG 1741
Db 1689 CCAGAACATTTACAATATATCTGGAAGTTGTACCAGAGAGATAAGTTATGAACTAATG 1748
QY 1742 GACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAATAGGATTTTACTTCTATT 1801
Db 1749 GACTTGAGATCTTGGCAATCTGCCCAAGGGGAGACACAAAATAGGATTTTACTTCTATT 1808
QY 1802 TTCTGAAAGCTAGAGATTTACAATCTTTGGTGATATAACAAA 1843
Db 1809 TTCTGAAAGCTAGAGATTTACAATCTTTGGTGATATAACAAA 1850

RESULT 8
AB040875
LOCUS
DEFINITION Homo sapiens hxCT mRNA for cystine/glutamate exchanger, complete cds.
ACCESSION AB040875
VERSION AB040875.1
KEYWORDS GI:13516845
SOURCE
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kim, J.Y., Chairoungdua, A., Cha, S.H., Segawa, H., Matsuo, H., Kim, D.K., Endou, H. and Kanai, Y.
TITLE Human cystine/glutamate exchanger: cDNA cloning and upregulation by oxidative stress in glioma cells
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2000)
AUTHORS Kanai, Y.
TITLE Direct Submission
JOURNAL Submitted (01-APR-2000) Yoshikatsu Kanai, Kyorin University School of Medicine, Department of Pharmacology and Toxicology; 6-20-2 Shinkawa, Mitaka, Tokyo 181-8611, Japan
(E-mail: ykanai@kyorin-u.ac.jp, Tel: +81-422-47-5511 (ex. 3453), Fax: +81-422-79-1321)

FEATURES
source
1. .2000
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hxCT"
/clone_lib="brain cDNA library"
/dev_stage="adult"
1. .2000
/gene="hxCT"
136..1620
/gene="hxCT"
/note="Na+-independent amino acid transporter transmembrane protein"
/codon_start=1
/product="cystine/glutamate exchanger"
/protein_id="BAB40574.1"
/db_xref="GI:13516846"
/translation="MVRKPVVSTISKGYLQGVNNGRLPSLGNKEPPQKQVQLKRY TLLRVSIIITGIIGAFIFSPKQLQNTSGVMSLTITWTCVLSLFGALSYAELGT TIKKSGHYTVILEVFGPLPAFVRVWVLLIIIRPAATAVISLAFGRYILEPFFIQCSI PELAKLTAVITVWVLSMSVWSARIOIFUTCKLTALITIIIVPGVQLIKGOT ONFKDTSGRDSSITRLPLAFYCYMAYAGWFYLNFTVEENPEKTIPLAICISMAI VTIGVLTNVAIFTITINAEULLLSNNAVATFSERLGNFSLAVPIFVALSCFSGMNG VFAVSRIFYASREGHILPMIHRKHTPLPAIVLHPLTMIPLFSGDLSLNLFL

Db 1325 CTCTACAGCTGTTATTGTTTGGACCCCTTTGACAAATGATAATGCTCTCTCTGGAGACC 1384
 QY 1382 TCACAGCTCTTTTGAATTTCCCTCAGTTTGGCCAGGTGGCTTTTATTTGGGCTGGCAGTTG 1441
 Db 1385 TCACAGCTCTTTTGAATTTCCCTCAGTTTGGCCAGGTGGCTTTTATTTGGGCTGGCAGTTG 1444
 QY 1442 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTCTTCAAGGTGCCAC 1501
 Db 1445 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTCTTCAAGGTGCCAC 1504
 QY 1502 TGTTCATCCAGCTTTGTTTTCCTTCACATGCTCTTCATGTTGGCCCTTTCCCTCTATT 1561
 Db 1505 TGTTCATCCAGCTTTGTTTTCCTTCACATGCTCTTCATGTTGGCCCTTTCCCTCTATT 1564
 QY 1562 CGGACCCATTAGTACAGGATTTGGCTTCTGTCATCACTCTGACTGGAGTCCCTGGCTATT 1621
 Db 1565 CGGACCCATTAGTACAGGATTTGGCTTCTGTCATCACTCTGACTGGAGTCCCTGGCTATT 1624
 QY 1622 ATCTCTTTTATATATGGCAAGAAACCCAGGTGGTTTGAATAATGTTCAGAGAAATAA 1681
 Db 1625 ATCTCTTTTATATATGGCAAGAAACCCAGGTGGTTTGAATAATGTTCAGAGAAATAA 1684
 QY 1682 CCAGACATTACAAATAATCTGGAAGTTGTACCAGAGAATAAGTTATGAACATAAG 1741
 Db 1685 CCAGACATTACAAATAATCTGGAAGTTGTACCAGAGAATAAGTTATGAACATAAG 1744
 QY 1742 GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGATTTTACTTCAAT 1801
 Db 1745 GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGATTTTACTTCAAT 1804
 QY 1802 TTCTGAAAGCTAGAGAAATTTACAATTTGGTGATTAACAAAGGAGTCAAGTTATTTTAT 1861
 Db 1805 TTCTGAAAGCTAGAGAAATTTACAATTTGGTGATTAACAAAGGAGTCAAGTTATTTTAT 1864
 QY 1862 TCATATATTTAGCATATTCGAATTAATTTCTAAGAAATTTAGTTATATACTCTATGACT 1921
 Db 1865 TCATATATTTAGCATATTCGAATTAATTTCTAAGAAATTTAGTTATATACTCTATGACT 1924
 QY 1922 TATAGAAAGTGAATATGCACTTATCTATGAGTCGCACAAATCTTCTGATACCT 1981
 Db 1925 TATAGAAAGTGAATATGCACTTATCTATGAGTCGCACAAATCTTCTGATACCT 1984
 QY 1982 ACCTATTTGGGTTAGGAGAAAGACTAGACAATTTACTATGTTGCTATCTTCAACATA 2041
 Db 1985 ACCTATTTGGGTTAGGAGAAAGACTAGACAATTTACTATGTTGCTATCTTCAACATA 2044
 QY 2042 TGTAGCAGCGCAAGAACCTTCAAAATGGAAGACTGAGATTTTCTGTATATATGGGTT 2101
 Db 2045 TGTAGCAGCGCAAGAACCTTCAAAATGGAAGACTGAGATTTTCTGTATATATGGGTT 2104
 QY 2102 TGTAAAGATGGTTTACACACTACAGATGCTATCTATGTTGAAAGTGTTCATTTCTGA 2161
 Db 2105 TGT-AAAGATGGTTTACACACTACAGATGCTATCTATGTTGAAAGTGTTCATTTCTGA 2163
 QY 2162 AAAAAACATACATCATGATTATGGCAAGAGAGAGAGGTA 2204
 Db 2164 AAAAAACATACATCATGATTATGGCAAGAGAGAGAGAA 2206

RESULT 6
 AF200708 1874 bp mRNA linear PRI 01-DEC-2000
 LOCUS Homo sapiens calcium channel blocker resistance protein CCBRL mRNA, complete cds.
 DEFINITION
 ACCESSION AF200708
 VERSION AF200708.1 GI:11493651
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1874)
 AUTHORS Conklin,D.S. and Beach,D.H.

TITLE CCBRL, novel CD98 light chain implicated in redox control and calcium signaling
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1874)
 AUTHORS Conklin,D.S. and Beach,D.H.
 TITLE Direct Submission
 JOURNAL Submitted (01-NOV-1999) Cold Spring Harbor Labs, 1 Bungtown Rd, Cold Spring Harbor, NY 11724, USA
 FEATURES
 Location/Qualifiers
 1..1874
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="fibrosarcoma HT1080"
 246..1751
 /note="X-c transporter; CD98 light chain"
 /codon_start=1
 /product="calcium channel blocker resistance protein CCBRL"
 /protein_id="AAG35592.1"
 /db_xref="GI:11493652"
 /translation="MVRKPVSTISKGYLQGNVGRPLSLGNKEPPGQEKVLRKV TLRGVSIITIGAGIFLSPKVLQNTGSGVMSLTITVCGVLSLFGALSVAELGT TIKSGHYIYIILEVFGPLPAFVRVWVELLIIRPATAVISLAFGRYILPFFIQCEI PELAKLITAVGITVMVNSVMSVARSIOELFTCKLTAILIILVPGVMQIKQGT QNFDAFSGRDSITRPLAFYTGMYAYAGWYLFNFTEVENPEKTIPLAICISMAI VTGYVLTVAYFTTINAEELLSNVAVTSERLGNFSLAVIFVSLGSCFSGMNG VFAYSLFYVASREGHLEILSMIHVRKHTPLPAVIVLHPLTMIPLFSGDLSDLLNPL SFARWLFGLAVAGLIYLRVKCPDMHRPKVPLFIPALFSTCLEFVALVSLYSDPST GIGEVLTGVVPAYLFIINDKPRFIMSEKTRFLQIILLEVVEDKL"
 BASE COUNT 490 a 407 c 422 g 555 t
 ORIGIN
 Query Match 68.9%; Score 1542.4; DB 9; Length 1874;
 Best Local Similarity 96.7%; Pred. No. 0;
 Matches 1607; Conservative 0; Mismatches 1; Indels 54; Gaps 1;
 QY 182 GAAAGCCCTGTTGTCTCCACCATCTCCAAAGGAGGTACTCGACGGAAATGTTAACGGGA 241
 Db 253 GAAAGCCCTGTTGTCTCCACCATCTCCAAAGGAGGTACTCGACGGAAATGTTAACGGGA 312
 QY 242 GGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGACGACGCTTTCAGGAAGAGAGACG 301
 Db 313 GGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGACG----- 350
 QY 302 CCTTTTCAGGAAGACGCGCTTTTCAGGAAGAGAAAGTGCAGCTGAAGAGAAAGTCA 361
 Db 351 -----GAGAAAGTGCAGCTGAAGAGAAAGTCA 378
 QY 362 CTTTACTGAGGGAGTCTCCATTATCATTTGGACCATCATTTGGACGAGAAATCTTCATCT 421
 Db 379 CTTTACTGAGGGAGTCTCCATTATCATTTGGACCATCATTTGGACGAGAAATCTTCATCT 438
 QY 422 CTCCTAAGGGCTGCTCCAGAACACGGCAGCGTGGGATGTCTCTGACCATCTGGACGG 481
 Db 439 CTCCTAAGGGCTGCTCCAGAACACGGCAGCGTGGGATGTCTCTGACCATCTGGACGG 498
 QY 482 TGTGTGGGTCCTGTCTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAATAA 541
 Db 499 TGTGTGGGTCCTGTCTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAATAA 558
 QY 542 AGAAATCTGGAGGTCATTACACATATATTTGGAAGTCTTTGGTCCATTACAGCTTTTG 601
 Db 559 AGAAATCTGGAGGTCATTACACATATATTTGGAAGTCTTTGGTCCATTACAGCTTTTG 618
 QY 602 TACGAGTCTGGTGGAACTCCCTCAATAAGCCCTGCAGCTACTCTCTGTATATCCCTGG 661
 Db 619 TACGAGTCTGGTGGAACTCCCTCAATAAGCCCTGCAGCTACTCTCTGTATATCCCTGG 678
 QY 662 CATTTGGAGCTACATTTCTGGAAACCATTTTATTCAATGTGAAATCCCTGAACCTGGGA 721
 Db 679 CATTTGGAGCTACATTTCTGGAAACCATTTTATTCAATGTGAAATCCCTGAACCTGGGA 738
 QY 722 TCAAGCTCATTAACAGTGTGGGCATAAAGTGTAGTCTGATGCTCCTAAATAGCATGAGTCA 781

```

Db 1854 TATAGAAAGTGAATATGCAGTTATCTATGATCGCACAAATCTTGAGTCTCTGATACCT 1913
QY 1982 ACCTATTGGGGTTAGGAGAAAGACTAGACAAATCTATGTTGGTCAATCTCTACACATA 2041
Db 1914 ACCTATTGGGGTTAGGAGAAAGACTAGACAAATCTATGTTGGTCAATCTCTACACATA 1973
QY 2042 TGTAGCAGCGGCAAGAACCTTCAAAATGAAGACTGAGATTTTCTGTTATATATGGTTT 2101
Db 1974 TGTAGCAGCGCAAGAACCTTCAAAATGAAGACTGAGATTTTCTGTTATATATGGTTT 2033
QY 2102 TGTAAAGATGGTTTACACACTACAGATGCTATATCTGTGAAAGTGTTTCAAAATCTGA 2161
Db 2034 TGTAAAGATGGTTTACACACTACAGATGCTATATCTGTGAAAGTGTTTCAAAATCTGA 2093
QY 2162 AAAAAAGCATACATCATGATTATGGCAAGAGAGAGAA 2200
Db 2094 AAAAAAGCATACATCATGATTATGGCAAAAAA 2132

RESULT 5
AF252872
LOCUS AF252872 2482 bp mRNA linear PRI 02-MAY-2001
DEFINITION Homo sapiens cystine/glutamate transporter xCT mRNA, complete cds.
ACCESSION AF252872
VERSION AF252872.1 GI:13924719
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2482)
AUTHORS Chancy,C.D., Kekuda,R., Wang,H., Huang,W., Prasad,P.D., Smith,S.B.
and Ganapathy,V.
TITLE Structure, Function and Regulation of Human Cystine/Glutamate
Transporter in Retinal Pigment Epithelial Cells
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2482)
AUTHORS Wang,H., Prasad,P.D. and Ganapathy,V.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2000) Biochemistry & Molecular Biology, Medical
College of Georgia, 1120 15th Street, Augusta, GA 30912, USA
FEATURES
Source
1..2482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
232..1737
/codon_start=1
/product="cystine/glutamate transporter xCT"
/protein_id="AAK4911.1"
/db_xref="GI:13924720"
/translation="MVRPVSSTISKGYLOGNVNGLRPLSLGNKEPPQEKVQLKRV
TLRGSIIITIGAGIFISPKGLQNTGVSMSLTITWYCVGLSLFGALSYAELGT
TIKSGSHYITVILEVFGPLPAFVRWVLLIIRPAATAVIAFLAFGRVILEPFFQCEI
PELAIKLTAVITVVMVLSMSYSWSARIQIFLTKLAILIIIVPGVMOLIKGOT
QNFKDAFGRSDSTIRLPLAFYGYMYAVAGFYLNFTVEENKEKPIPAICISMAI
VTIGVNLNVAFTTINAEELLLSNVAVTSERLGNFSLAVIFVALSCFSGMNG
VFAVSRFYVASRSHLPPEILSMIHVRKHPPLPAVIVLHPLTMILPFGDLSLLNEL
SFARWLTIGLAVAGLILYRKYCPDMRFPKPLPFPALFSEFCFLMFWALSYPDFST
GIGFYITLIGVAPYVILFIWDKPKPRFMSEKIRTRLIQIILEVVPEDKL"
BASE COUNT 735 a 484 c 521 g 742 t
ORIGIN
Query Match 84.3%; Score 1888.2; DB 9; Length 2482;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1965; Conservative 0; Mismatches 3; Indels 55; Gaps 2;

QY 182 GAAAGCCTGTGTGTCACCATCTCAAGAGAGGTTACTGCGAGGAAATGTTACGGGA 241
Db 239 GAAAGCCTGTGTGTCACCATCTCAAGAGAGGTTACTGCGAGGAAATGTTACGGGA 298
QY 242 GCGTCCCTTCCCTGGGCAACAGAGGCCACCTGGGCGAGCGCTTTTCAGGAAGAGACG 301

```

```

Db 299 GCGTCCCTTCCCTGGGCAACAGAGGCCACCTGGGCGAG----- 336
QY 302 CCTTTTCAGGAAGAGAGAGCGCCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAGTCA 361
Db 337 -----GAGAAAGTGCAGCTGAAGAGGAAGTCA 364
QY 362 CTTTACTGAGGGGAGTCTCCATTTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATCT 421
Db 365 CTTTACTGAGGGGAGTCTCCATTTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCATCT 424
QY 422 CTCCTTAAGGGGCTGCTCCAGAACACGGGCGAGCGTGGGCATCTCTGACCATCTCTGGACGG 481
Db 425 CTCCTTAAGGGGCTGCTCCAGAACACGGGCGAGCGTGGGCATCTCTGACCATCTCTGGACGG 484
QY 482 TGTGTGGGGTCTCTGTCACATTTATTTGGAGCTTTTGTCTTATGCTGAATGGGAAACAATAAA 541
Db 485 TATGTGGGGTCTCTGTCACATTTATTTGGAGCTTTTGTCTTATGCTGAATGGGAAACAATAAA 544
QY 542 AGAAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTTGGTCCATTACCAGCTTTTG 601
Db 545 AGAAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTTGGTCCATTACCAGCTTTTG 604
QY 602 TACGAGTCTGGGTGGAACTCTCTCATATATACGCCCTGCGAGCTACTGCTGTGATATCCCTGG 661
Db 605 TACGAGTCTGGGTGGAACTCTCTCATATATACGCCCTGCGAGCTACTGCTGTGATATCCCTGG 664
QY 662 CATTTGGAGCGTACATTTCTGGAACCAATTTTTTATTTCAATGGAATCCCTGAACTGCGA 721
Db 665 CATTTGGAGCGTACATTTCTGGAACCAATTTTTTATTTCAATGGAATCCCTGAACTGCGA 724
QY 722 TCAAGCTCATTTACAGCTGTGGGCATAACTGTAGTGATGGTCTCTAAATAGCATGAGTGTCA 781
Db 725 TCAAGCTCATTTACAGCTGTGGGCATAACTGTAGTGATGGTCTCTAAATAGCATGAGTGTCA 784
QY 782 GCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTCAAGCTCACAGCAATTTCTGATAA 841
Db 785 GCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTCAAGCTCACAGCAATTTCTGATAA 844
QY 842 TTATAGTCCCTGGAGTATGATGAGCTAATTAAGGTCAACGAGCACTTTAAAGACGCGCT 901
Db 845 TTATAGTCCCTGGAGTATGATGAGCTAATTAAGGTCAACGAGCACTTTAAAGACGCGCT 904
QY 902 TTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATATATGGAATGTATG 961
Db 905 TTTCAGGAAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATATATGGAATGTATG 964
QY 962 CATATGCTGGCTGCTTTTACCTCAACTTTGTTTACTGAAGAAGTAGAAAAACCTCGAAAAA 1021
Db 965 CATATGCTGGCTGCTTTTACCTCAACTTTGTTTACTGAAGAAGTAGAAAAACCTCGAAAAA 1024
QY 1022 CCATTCCCTTGGCAATATGATATATCCATGGCCATTTGTCACCATTTGGCTATGTGCTGACAA 1081
Db 1025 CCATTCCCTTGGCAATATGATATATCCATGGCCATTTGTCACCATTTGGCTATGTGCTGACAA 1084
QY 1082 ATGTGGCTACTTTACGACCATTAATCTCAGGAGCTGCTGCTTTCAATTCAGTGGCAG 1141
Db 1085 ATGTGGCTACTTTACGACCATTAATCTCAGGAGGCTGCTGCTTTCAATTCAGTGGCAG 1144
QY 1142 TGACCTTTTCTGAGCGGCTACTGGGAATTTCTCATAGCAGTTCCCATCTTTGTTGCC 1201
Db 1145 TGACCTTTTCTGAGCGGCTACTGGGAATTTCTCATAGCAGTTCCCATCTTTGTTGCC 1204
QY 1202 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTTCTATGTTG 1261
Db 1205 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTTCTATGTTG 1264
QY 1262 GGTCTCAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCCGCAAGCACACTC 1321
Db 1265 GGTCTCAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCCGCAAGCACACTC 1324
QY 1322 CTCCTACAGCTGTTATTTGTTTGGCACCCCTTTGACAATGATAATGCTCTTCTCTGGAGACC 1381

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:20026 IMAGE:4562994"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
161..1666
CDS
/codon_start=1
/product="Similar to solute carrier family 7, (cationic
amino acid transporter, y+ system) member 11"
/protein_id="AAH12087.1"
/db_xref="GI:15082352"
/translation="MVRKPVVSTISKGYLQGVNGLPSLGNKEPPGKVKLRKV
TLKRGVSIITIGAGIFISPKGVQLNTGVSGLNITWVCGVLSLFGALSYAELGT
TIKSGGHYIIEVFGPLPAFVPLVVELLIIRPAATVAVISLAFGRYLDPPFIQCEI
PELAIKITATVIVVLMNSVMSARQIFLTKLTAIILIIIPVGMQIKQOT
QNFKDAFSGRDSITRPLAFYGMAYAGWFLNFVEENPEKTIPLAIGISMAI
VTIGVLTNVAFTTINAEILLNSAVATFBRLLGNFSLAVPIFVLSFCDSMNGG
VFVSRFLYVAVSRGHLPEILSMITHVSKHTPLPAVIVLHPLTMILFSGDLDSLLNPL
SFARWLFIGLAVAGLIYLRKCPDMRPPKVPFLFIPALFSFTCLFWVALSLYSDPST
GIGVITLTGVPAYLFIILWKKPRFRIMSEKIITRLQIILEVVPEDKL"
BASE COUNT 599 a 443 c 452 g 661 t
ORIGIN
Query Match 84.3%; Score 1888.2; DB 9; Length 2155;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1957; Conservative 0; Mismatches 8; Indels 54; Gaps 1;
QY 182 GAAAGCGCTGTTGTGCCACCATCTCCAAAGGAGGTTACCTGCAGGAAATGTTAACGGGA 241
DB 168 GAAAGCGCTGTTGTGCCACCATCTCCAAAGGAGGTTACCTGCAGGAAATGTTAACGGGA 227
QY 242 GCGTCCCTTCCCTGGGCAACAGGAGCCACCTGGCGCAGCAGCCCTTTTCAGGAAGAGCG 301
DB 228 GCGTCCCTTCCCTGGGCAACAGGAGCCACCTGGCGAG----- 265
QY 302 CCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTCA 361
DB 266 -----GAGAAGTGCAGCTGAAGAGAAAGTCA 293
QY 362 CTTTACTAGGGAGTCTCCATATCATATGGCACCATCATATGGAGCAGGAATCTTCATCT 421
DB 294 CTTTACTAGGGAGTCTCCATATCATATGGCACCATCATATGGAGCAGGAATCTTCATCT 353
QY 422 CTCCTAAGGCGTGTCCAGAACACGGCAGCGTGGCATGTCTCTGACCATCTGACGG 481
DB 354 CTCCTAAGGCGTGTCCAGAACACGGCAGCGTGGCATGTCTCTGACCATCTGACGG 413
QY 482 TGTGTGGGTCTGTGTCACATATTTGGAGCTTTGTCTTATGCTGAATGGGAACAATAA 541
DB 414 TGTGTGGGTCTGTGTCACATATTTGGAGCTTTGTCTTATGCTGAATGGGAACAATAA 473
QY 542 AGAAATCTGGAGTCAATACATATATTTTGGAGCTTTTGGTCCATACCACTTTTG 601
DB 474 AGAAATCTGGAGTCAATACATATATTTTGGAGCTTTTGGTCCATACCACTTTTG 533
QY 602 TACGAGTCTGGGGAACCTCTCATATACGCCCTGACGCTACTGTGTGATATCCCTGG 661
DB 534 TACGAGTCTGGGGAACCTCTCATATACGCCCTGACGCTACTGTGTGATATCCCTGG 593
QY 662 CATTTGGAGCTACATCTTGGAAACCATTTTTTATTCATATGTGAATTCCTGAACCTTGGGA 721
DB 594 CATTTGGAGCTACATCTTGGAAACCATTTTTTATTCATATGTGAATTCCTGAACCTTGGGA 653
QY 722 TCAAGCTCATACAGTCTGGGCAATCACTGTAGTGTGCTTAATACATGAGTGCA 781
DB 654 TCAAGCTCATACAGTCTGGGCAATCACTGTAGTGTGCTTAATACATGAGTGCA 713
QY 782 GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTTCAGCAATCTGATAA 841
DB 714 GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTTCAGCAATCTGATAA 773

```

```

QY 842 TTATAGTCCCTGGAGTTATGACAGTAATTAAGGTCAAAACGACGAACTTTAAAGCGCT 901
DB 774 TTATAGTCCCTGGAGTTATGACAGTAATTAAGGTCAAAACGACGAACTTTAAAGCGCT 833
QY 902 TTTACGAGAGAGATTCAAGTATTAACGGTGTGCCACTGCTTTTATATGGAATGATG 961
DB 834 TTTACGAGAGAGATTCAAGTATTAACGGTGTGCCACTGCTTTTATATGGAATGATG 893
QY 962 CATATGCTGCTGCTTTTACCTCAACTTTTACTCAAGAAAGTAGAAAAACCTTGAAGAAA 1021
DB 894 CATATGCTGCTGCTTTTACCTCAACTTTTACTCAAGAAAGTAGAAAAACCTTGAAGAAA 953
QY 1022 CCATTTCCCTTGGCAATATGATATATCCATGCGCATTTGTCACCATTTGCTGCTGACAA 1081
DB 954 CCATTTCCCTTGGCAATATGATATATCCATGCGCATTTGTCACCATTTGCTGCTGACAA 1013
QY 1082 ATGTGGCTACTTTAGACCATTAATGCTGAGAGAGCTGCTGCTTTCAATATGAGTGGCAG 1141
DB 1014 ATGTGGCTACTTTAGACCATTAATGCTGAGAGAGCTGCTGCTTTCAATATGAGTGGCAG 1073
QY 1142 TGACCTTTCTGAGCGCTACTCGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC 1201
DB 1074 TGACCTTTCTGAGCGCTACTCGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC 1133
QY 1202 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTTATCTATGTTG 1261
DB 1134 TCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTTATCTATGTTG 1193
QY 1262 GGTCTCAGAGGGTCACTTTCCAGAAATCTCTCCATGATTGATTCGCGCAACACACATC 1321
DB 1194 GGTCTCAGAGGGTCACTTTCCAGAAATCTCTCCATGATTGATTCGCGCAACACACATC 1253
QY 1322 CTCTACAGCTGTTTATGTTTTCACCCCTTTGACAAATGATAATGCTCTCTCGGAGACC 1381
DB 1254 CTCTACAGCTGTTTATGTTTTCACCCCTTTGACAAATGATAATGCTCTCTCGGAGACC 1313
QY 1382 TCGACAGCTTTTGAATTTCTCAGTTTTCGAGGCTGGCTTTTATTTGGGCTGGCAGTTG 1441
DB 1314 TCGACAGCTTTTGAATTTCTCAGTTTTCGAGGCTGGCTTTTATTTGGGCTGGCAGTTG 1373
QY 1442 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGCTCTTCAAGTGGCCAC 1501
DB 1374 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGCTCTTCAAGTGGCCAC 1433
QY 1502 TGTTCATCCAGCTTTGTTTTCCTTCACATGCTCTTCATGGTGGCTTTCCCTCTATT 1561
DB 1434 TGTTCATCCAGCTTTGTTTTCCTTCACATGCTCTTCATGGTGGCTTTCCCTCTATT 1493
QY 1562 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTCGCTATT 1621
DB 1494 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACTGGAGTCCCTCGCTATT 1553
QY 1622 ATCTCTTTATTTATTTGGGCAAGAAACCCAGGTGGTGTAGAAATATGTCAGAGAAAAATA 1681
DB 1554 ATCTCTTTATTTATTTGGGCAAGAAACCCAGGTGGTGTAGAAATATGTCAGAGAAAAATA 1613
QY 1682 CCAGAACATTTACAAATTAATCTGGAAGTTGTACCAGAGAGATTAAGTTATGAACATAG 1741
DB 1614 CCAGAACATTTACAAATTAATCTGGAAGTTGTACCAGAGAGATTAAGTTATGAACATAG 1673
QY 1742 GACTTGAGATCTTGGCATCTGCCCAAGGGGAGACACAAAAATAGGATTTTACTTCTATT 1801
DB 1674 GACTTGAGATCTTGGCATCTGCCCAAGGGGAGACACAAAAATAGGATTTTACTTCTATT 1733
QY 1802 TTCTGAAAGTCTAGAGAAATACAACTTTTGGTGATAAAACAAAGAGTCAAGTATTTTAT 1861
DB 1734 TTCTGAAAGTCTAGAGAAATACAACTTTTGGTGATAAAACAAAGAGTCAAGTATTTTAT 1793
QY 1862 TCATATATTTTATGATATTCGAACATTAATTTCTAAGAAATTTAGTTATTAACCTCTATGAT 1921
DB 1794 TCATATATTTTATGATATTCGAACATTAATTTCTAAGAAATTTAGTTATTAACCTCTATGAT 1853
QY 1922 TATAGAAAGTGAATATGAGTTATTTCTATGATGTCGCAACATTTCTTGAGTCTCTGATACCT 1981

```

QY 722 TCAAGCTCATTACAGCTGTGGGCAACATGTAGTGTGCTCTAAATAGCATGAGTGCA 781
 Db 494 TCAAGCTCATTACAGCTGTGGGCAACATGTAGTGTGCTCTAAATAGCATGAGTGCA 553
 QY 782 GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCAGCAATTCGTGATA 841
 Db 554 GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCAGCAATTCGTGATA 613
 QY 842 TTATAGTCCCTGGAGTATGAGCTAATTAAGGTTCAACGCGAGCACTTAAAGACGGCT 901
 Db 614 TTATAGTCCCTGGAGTATGAGCTAATTAAGGTTCAACGCGAGCACTTAAAGACGGCT 673
 QY 902 TTTCAGGAAGAGATCAAGTATTAACCGGTTGCCACTGGCTTTTATTAATGGAATGTATG 961
 Db 674 TTTCAGGAAGAGATCAAGTATTAACCGGTTGCCACTGGCTTTTATTAATGGAATGTATG 733
 QY 962 CATATGCTGGCTGGTTTACCTCAACTTTTACTGAAGAAGTAGAAAACCCCTGAAAAA 1021
 Db 734 CATATGCTGGCTGGTTTACCTCAACTTTTACTGAAGAAGTAGAAAACCCCTGAAAAA 793
 QY 1022 CCATTCCTCCCTTGAATATATATCCATGGCCATTTGTCACCATTTGGCTATGTGCTGACAA 1081
 Db 794 CCATTCCTCCCTTGAATATATATCCATGGCCATTTGTCACCATTTGGCTATGTGCTGACAA 853
 QY 1082 ATGTGGCTTACTTTAGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGACATGCGAG 1141
 Db 854 ATGTGGCTTACTTTAGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGACATGCGAG 913
 QY 1142 TGACCTTTTCTGAGCGGCTACTGGCAATTTCTCAATAGCAGTTCGGATCTTTGTTGCC 1201
 Db 914 TGACCTTTTCTGAGCGGCTACTGGCAATTTCTCAATAGCAGTTCGGATCTTTGTTGCC 973
 QY 1202 TCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTTCTATGTTG 1261
 Db 974 TCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGTCTCCAGGTTATTTCTATGTTG 1033
 QY 1262 CGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCATGATTCATGTCGCAAGCACACTC 1321
 Db 1034 CGTCTCGAGAGGGTCACTTCCAGAAATCCTCTCATGATTCATGTCGCAAGCACACTC 1093
 QY 1322 CTTACACAGCTTTATTTGTTGACCCCTTTGACAAATCATATGCTCTCTCTCGAGACC 1381
 Db 1094 CTTACACAGCTTTATTTGTTGACCCCTTTGACAAATCATATGCTCTCTCTCGAGACC 1153
 QY 1382 TCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGCTTTTATTTGGGCTGGCAGTTG 1441
 Db 1154 TCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGCTTTTATTTGGGCTGGCAGTTG 1213
 QY 1442 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTTCAAGTGGCAC 1501
 Db 1214 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTTCAAGTGGCAC 1273
 QY 1502 TGTTCATCCAGCTTTGTTTCTTCATCATGCTCTTCATGTTGGCCTTTCCCTCTATT 1561
 Db 1274 TGTTCATCCAGCTTTGTTTCTTCATCATGCTCTTCATGTTGGCCTTTCCCTCTATT 1333
 QY 1562 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACCTGGAGTCCCTCGGTATT 1621
 Db 1334 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACCTGGAGTCCCTCGGTATT 1393
 QY 1622 ATCTCTTTTATTTATGGGACAAAGAACCCAGGTGGTTTAGAATATGTCAGAGAAATAA 1681
 Db 1394 ATCTCTTTTATTTATGGGACAAAGAACCCAGGTGGTTTAGAATATGTCAGAGAAATAA 1453
 QY 1682 CCAGAACATTACAATAATCTGGAAGTTGTACCAGAGAGAGATAAGTTATGAACATAATG 1741
 Db 1454 CCAGAACATTACAATAATCTGGAAGTTGTACCAGAGAGAGATAAGTTATGAACATAATG 1513
 QY 1742 GACTTGAGATCTGGCAATCGCCCAAGGGGAGACACAAAATAGGGATTTTTACTTTCATT 1801
 Db 1514 GACTTGAGATCTGGCAATCTGCCCAAGGGGAGACACAAAATAGGGATTTTTACTTTCATT 1573
 QY 1802 TTCTGAAAGCTAGAGAATTACAACTTTGGTGATAAACAAGAGTCAGTTATTTTAT 1861

Db 1574 TTCTGAAGCTAGAGATTCACAACTTTGGTGATAAACAAGAGTCAGTTATTTTAT 1633
 QY 1862 TCATATATTTTACGATATTCGAACATAATTTCTTAAGAAATTTAGTTATAACTCTATGTAGT 1921
 Db 1634 TCATATATTTTACGATATTCGAACATAATTTCTTAAGAAATTTAGTTATAACTCTATGTAGT 1693
 QY 1922 TATAGAAGTGAATATGAGCTTATTTCTAGTGGCACAATTTCTTGAAGTCTCTGATACCT 1981
 Db 1694 TATAGAAGTGAATATGAGCTTATTTCTAGTGGCACAATTTCTTGAAGTCTCTGATACCT 1753
 QY 1982 ACCTATTGGGGTTAGGAGAAAGACTAGACAAATTTACTATGTTGGTCACTTCTCAACATA 2041
 Db 1754 ACCTATTGGGGTTAGGAGAAAGACTAGACAAATTTACTATGTTGGTCACTTCTCAACATA 1813
 QY 2042 TGTAGCACGGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT 2101
 Db 1814 TGTAGCACGGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT 1873
 QY 2102 TGTAAAGATGTTTTACACACTACAGATGCTATACTGTGAAAAGTGTTCATTTCTGA 2161
 Db 1874 TGTAAAGATGTTTTACACACTACAGATGCTATACTGTGAAAAGTGTTCATTTCTGA 1933
 QY 2162 AAAAAACATACATCATGATTATGGCAAGAGGAGAGAGGTA 2204
 Db 1934 AAAAAACATACATCATGATTATGGCAAGAGGAGAGAGAA 1976

RESULT 4
 BC012087
 LOCUS
 DEFINITION
 Homo sapiens, similar to solute carrier family 7, (cationic amino acid transporter, y+ system) member 11, clone MGC:20026
 IMAGE:4562994, mRNA, complete cds.
 ACCESSION
 BC012087
 VERSION
 BC012087.1 GI:15082351
 KEYWORDS
 MGC.
 SOURCE
 Homo sapiens.
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 2155)
 AUTHORS
 Strausberg, R.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (02-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcqsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 29 Row: 9 Column: 1
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5668544.
 Location/Qualifiers
 1. .2155

FEATURES
 source

Db 1327 CGACAGTCCTTTGAATTCCTCAGTTTCCAGTGGCTTTTATTTGGCTGGCAGTTCC 1386
 QY 1443 TGGGCTGATTATCTTCGATACAAATGCCAGATATGCATCGCTCTTCAAGTGCCACT 1502
 Db 1387 TGGGCTGATTATCTTCGATACAAATGCCAGATATGCATCGCTCTTCAAGTGCCACT 1446
 QY 1503 GTTCATCCCAAGCTTTGTTTCCCTTCACATGCGCTTCCTCATGCTTTCCTCTCATTC 1562
 Db 1447 GTTCATCCCAAGCTTTGTTTCCCTTCACATGCGCTTCCTCATGCTTTCCTCTCATTC 1506
 QY 1563 GGACCCATTTAGTACAGGATTTGGCTTCGTCATCATCTGCTGAGTCCCTGCGTATTA 1622
 Db 1507 GGACCCATTTAGTACAGGATTTGGCTTCGTCATCATCTGCTGAGTCCCTGCGTATTA 1566
 QY 1623 TCTCTTTATTTATATGGACAAAGAACCCAGTGGTTTGAATTAATCTCAGAGAAATAAC 1682
 Db 1567 TCTCTTTATTTATATGGACAAAGAACCCAGTGGTTTGAATTAATCTCAGAGAAATAAC 1626
 QY 1683 CAGAACATTTACAAATATCTGGAAGTTGTACCAGAGAGATAAGTTATGAACATAATGG 1742
 Db 1627 CAGAACATTTACAAATATCTGGAAGTTGTACCAGAGAGATAAGTTATGAACATAATGG 1686
 QY 1743 ACTTGAGATCTTGGCATCTGCCAAGGGGAGACACAAATAGGATTTTACTTCATTT 1802
 Db 1687 ACTTGAGATCTTGGCATCTGCCAAGGGGAGACACAAATAGGATTTTACTTCATTT 1746
 QY 1803 TCTGAAAGTCTAGAGAAATACAACTTTGGTGATAAAACAAAGAGTCAGTTATTTTATT 1862
 Db 1747 TCTGAAAGTCTAGAGAAATACAACTTTGGTGATAAAACAAAGAGTCAGTTATTTTATT 1806
 QY 1863 CATATATTTTACGATATTCGAACATAATTTCTAAGAAATTTAGTTATTAACCTATGATTT 1922
 Db 1807 CATATATTTTACGATATTCGAACATAATTTCTAAGAAATTTAGTTATTAACCTATGATTT 1866
 QY 1923 ATAGAAAGTGAATATCGATTTATCTATGATGCGCAATTTCTGATCTCTGATACCTA 1982
 Db 1867 ATAGAAAGTGAATATCGATTTATCTATGATGCGCAATTTCTGATCTCTGATACCTA 1926
 QY 1983 CCTATTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTTGTCATCTCTACACATAT 2042
 Db 1927 CCTATTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTTGTCATCTCTACACATAT 1986
 QY 2043 GTTAGCAGCGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGCGTTT 2102
 Db 1987 GTTAGCAGCGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGCGTTT 2046
 QY 2103 GTAAAGATGTTTACACACTACAGATGCTATCTACTGTGAAAGTGTCTTCAATCTGAA 2162
 Db 2047 GTAAAGATGTTTACACACTACAGATGCTATCTACTGTGAAAGTGTCTTCAATCTGAA 2106
 QY 2163 AAAAGCATACATCATGATTATGGCAAGAGAGAGAGAGTA 2204
 Db 2107 AAAAGCATACATCATGATTATGGCAAGAGAGAGAGAGTA 2148

RESULT 3
 LOCUS HSA277882 3144 bp mRNA linear Pri 04-JAN-2002
 DEFINITION Homo sapiens mRNA for cystine/glutamate transporter (XCT gene).
 ACCESSION AJ277882
 VERSION AJ277882.1 GI:18073361
 KEYWORDS cystine/glutamate transporter; XCT gene.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Borsani,G., Manzoni,M., Palacin,M., Pineda,M. and Gasol,E.
 Unpublished
 JOURNAL 2 (bases 1 to 3144)
 REFERENCE
 AUTHORS Bassi,M.T.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-2000) Bassi M.T., Telethon Institute of Genetics

FEATURES
 Location/Qualifiers
 1. 3144
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /map="4q28-q32"
 /cell_line="NT2, undifferentiated teratocarcinoma cell line"
 gene
 1. 3128
 /gene="XCT"
 CDS
 1. 1506
 /gene="XCT"
 /codon_start=1
 /product="cystine/glutamate transporter"
 /protein_id="CAC81905.1"
 /db_xref="GI:18073362"
 /translation="MVRKPVVSTISKGYLQGNVGRPLPSLGNKEPPGQEKVQLRKRY
 TLLRGVSIITIGIAGIFIPKGVQNTSGVMSLTIWTCVGLSFLGALSYAELGT
 TIKSGHYTYILEVFGPLPAFVWVVELLIIRPAATAVISLAFRVLIPFFIQCBI
 PELAKLITAVITVWVMSVMSARQIFLTCKLTALIIIVPGVQLIKGOT
 ONKDAFSGRDSITRLPLAFYVGMAYAGWYLFNFTVEENPEKTIPIAICISNAI
 VTIGYLTNVAFTTINABELLSNAVATFSEBRLGNFSLAVPIFVALSCFSGMNG
 VFVSRFLYVASREHLPILSMIHRRKHTPLPAVIVLHPLTMIPLMFGDLSLLNFL
 SFARWLFYLAVALGLIYLYKCPDMHPRFKVPLFIPALFSTCLFMVALSLYSDPFSF
 GIGVITLTGVPAYVLFILWKKPRWFRIMSEKITRTLQIILEVVPEDKL"
 repeat_region 2857..3127
 /rpt_family="Sine/Alu"
 polyA_site 3128
 /gene="XCT"
 BASE COUNT 912 a 602 c 649 g 981 t
 ORIGIN
 Query Match 84.9%; Score 1900.2; DB 9; Length 3144;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 1966; Conservative 0; Mismatches 3; Indels 54; Gaps 1;
 QY 182 GAAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTACCTGCAGGAAATGTTAACGGGA 241
 Db 8 GAAAGCCTGTTGTGCCACCATCTCCAAAGGAGGTACCTGCAGGAAATGTTAACGGGA 67
 QY 242 GCGTCCCTCCCTGGGCAACAGGAGCCACCTGGCGCAGACGCCCTTTTTCAGGAAGAGACG 301
 Db 68 GCGTCCCTCCCTGGGCAACAGGAGCCACCTGGCGAG----- 105
 QY 302 CCTTTTCAGGAAGACAGCCCTTTTTCAGGAAGAGAGAAAGTCA 361
 Db 106 -----GAGAAAGTCAGCTGAAGAGAAAGTCA 133
 QY 362 CTTTACTGAGGGAGTCTCCATTTATTTGGCACCATCATTTGGAGCAGGAATCTTCATCT 421
 Db 134 CTTTACTGAGGGAGTCTCCATTTATTTGGCACCATCATTTGGAGCAGGAATCTTCATCT 193
 QY 422 CTCTTAAGGGCTGCTCCAGAACACGGGAGCGGTGGGATGTCTCTGACCATCTGGACGG 481
 Db 194 CTCTTAAGGGCTGCTCCAGAACACGGGAGCGGTGGGATGTCTCTGACCATCTGGACGG 253
 QY 482 TGTGTGGGCTCTGCACACTATTTGGAGCTTTGCTTTATCTGAAATGGGAACAACATAA 541
 Db 254 TGTGTGGGCTCTGCACACTATTTGGAGCTTTGCTTTATCTGAAATGGGAACAACATAA 313
 QY 542 AGAAATCTGGAGTCTATACACATATTTTGGAAAGTCTTTGGTCCATTACCAAGCTTTTG 601
 Db 314 AGAAATCTGGAGTCTATACACATATTTTGGAAAGTCTTTGGTCCATTACCAAGCTTTTG 373
 QY 602 TACGAGTCTGGTGAACCTCTCATAAATACGCCCTGGAGCTACTGCTGTGATATCCCTGG 661
 Db 374 TACGAGTCTGGTGAACCTCTCATAAATACGCCCTGGAGCTACTGCTGTGATATCCCTGG 433
 QY 662 CATTTGGAGCTACATTTCTGGAACCATTTTTTATTCAATGTGAAATCCCTCAACTTGGGA 721
 Db 434 CATTTGGAGCTACATTTCTGGAACCATTTTTTATTCAATGTGAAATCCCTCAACTTGGGA 493

QY 1981 TACCTATTGGGTTAGGAGAAAGACTAGACAATTACTATGTGGTCAATCTCTACAACAT 2040
Db 1981 TACCTATTGGGTTAGGAGAAAGACTAGACAATTACTATGTGGTCAATCTCTACAACAT 2040
QY 2041 ATGTTAGCAGCGCAAGAACCTTCAAAATGAAGACTGAGATTTTCTGTATATATGGGTT 2100
Db 2041 ATGTTAGCAGCGCAAGAACCTTCAAAATGAAGACTGAGATTTTCTGTATATATGGGTT 2100
QY 2101 TTGTAAGATGCTTTTACACACTACAGATGCTATATCTGTGAAAGTGTCTTCAATTCG 2160
Db 2101 TTGTAAGATGCTTTTACACACTACAGATGCTATATCTGTGAAAGTGTCTTCAATTCG 2160
QY 2161 AAAAAAGCATACATCATGATTATGCGAAAGAGAGAGAAAGTGTGTTTCAATTCG 2220
Db 2161 AAAAAAGCATACATCATGATTATGCGAAAGAGAGAGAAAGTGTGTTTCAATTCG 2220
QY 2221 TATTAATAAAAAA 2239
Db 2221 TATTAATAAAAAA 2239

RESULT 2
AX321910
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT
ORIGIN

Patent: WO 0172295-A 441 04-OCT-2001;
CORIXA CORPORATION (US)
1. 5981
/organism="Homo sapiens"
/db_xref="taxon:9606"
1858 a 1029 c 1098 g 1996 t

Query Match 92.9%; Score 2080.8; DB 6; Length 5981;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 2146; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY 3 AGGTTGAAGTGAAGAGATCATGCCAGCCCTGGGTGACAGTGAAGTCTGTCTCAACAG 62
Db 1 AGGTTGAAGTGAAGAGATCATGCCAGCCCTGGGTGACAGTGAAGTCTGTCTCAACAG 60
QY 63 AATTAGGAAAGAAAGAAAGAAAGAGAGAGAGAAATTCAGGCCAATGTGSCA 122
Db 61 AATTAGGAAAGAAAGAAAGAAAGAGAGAGAGAAATTCAGGCCAATGTGSCA 120
QY 123 TAGATTTATCATATCTCGGATTTTGTGATTCCTTTTCTCATCATCTGATTCAGG 182
Db 121 TAGATTTATCATATCTCGGATTTTGTGATTCCTTTTCTCATCATCTGATTCAGG 180
QY 183 AAAGCCTGTTGTCTCCACATCTCCAAAGAGGTTTACCTGAGGAAATGTATACGGAG 242
Db 181 AAAGCCTGTTGTCTCCACATCTCCAAAGAGGTTTACCTGAGGAAATGTATACGGAG 240
QY 243 GCTGCTTCCCTGGGCAACAAAGAGGCCACCTGGGCGAGACGCCCTTTTCAGGAAGAGAGC 302
Db 241 GCTGCTTCCCTGGGCAACAAAGAGGCCACCTGGGCGAG----- 277
QY 303 CTTTTAGGAAGAGAGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTCAC 362

Db 278 -----GAGAAAGTGCAGCTGAAGAGGAAAGTCAC 306
QY 363 TTTACTAGGGGAGTCTCCATTTATCATTTGACACCATTCATTTGGAGCAGGAATCTTCATCTC 422
Db 307 TTTACTAGGGGAGTCTCCATTTATCATTTGACACCATTCATTTGGAGCAGGAATCTTCATCTC 366
QY 423 TCCCTAAGGCGTGTCTCCAGAACACGGCAGCGTGGGATGCTCTGACCATCTGGAGGTT 482
Db 367 TCCTAAGGCGTGTCTCCAGAACACGGCAGCGTGGGATGCTCTGACCATCTGGAGGTT 426
QY 483 GTGTGGGCTCTCTCACTATTGAGGCTTTGTCTTATGCTGAATTTGGGAACAACATAAAA 542
Db 427 GTGTGGGCTCTCTCACTATTGAGGCTTTGTCTTATGCTGAATTTGGGAACAACATAAAA 486
QY 543 GAAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTACCACTTTTGT 602
Db 487 GAAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTACCACTTTTGT 546
QY 603 ACGAGTCTGGTGGAGTCTCTCATATACGCCCTGGAGCTACTGCTGTGATATCCCTGGC 662
Db 547 ACGAGTCTGGTGGAGTCTCTCATATACGCCCTGGAGCTACTGCTGTGATATCCCTGGC 606
QY 663 ATTTGGACGCTACATCTTGGAAACCATTTTATTCAATGTGAATCCCTGAACCTTGGCAT 722
Db 607 ATTTGGACGCTACATCTTGGAAACCATTTTATTCAATGTGAATCCCTGAACCTTGGCAT 666
QY 723 CAAGCTCATTTACAGCTGTGGGCATAACTGATGATGGTCTAAATAGCATAGTGTCTAG 782
Db 667 CAAGCTCATTTACAGCTGTGGGCATAACTGATGATGGTCTAAATAGCATAGTGTCTAG 726
QY 783 CTGAGGCGCCGATCCAGATTTTCTTAACCTTTTCAAGCTTCACAGCTCACAGCAATCTGATAAT 842
Db 727 CTGAGGCGCCGATCCAGATTTTCTTAACCTTTTCAAGCTTCACAGCTCACAGCAATCTGATAAT 786
QY 843 TATAGTCCCTGGAGTATGAGCTAAATTAAGGTCAAGGAGAGAAACCTTAAAGACGCCCTT 902
Db 787 TATAGTCCCTGGAGTATGAGCTAAATTAAGGTCAAGGAGAGAAACCTTAAAGACGCCCTT 846
QY 903 TCCAGGAAGAGATCAAGATTTACGGGTTTGCACCTGGCTTTTATTTATGGAATGTATGC 962
Db 847 TCCAGGAAGAGATCAAGATTTACGGGTTTGCACCTGGCTTTTATTTATGGAATGTATGC 906
QY 963 ATATGCTGGCTGCTTTTACCTCAACTTTTGTACTGAAGAGTAGAAACCTTGAAGAAAC 1022
Db 907 ATATGCTGGCTGCTTTTACCTCAACTTTTGTACTGAAGAGTAGAAACCTTGAAGAAAC 966
QY 1023 CATTCCTCTGCAATATGATATCCATGGCCATTGTCAACATTTGGCTATGTGTCAGAAA 1082
Db 967 CATTCCTCTGCAATATGATATCCATGGCCATTGTCAACATTTGGCTATGTGTCAGAAA 1026
QY 1083 TGTGGCTTACTTTACAGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAAGTGGCAGT 1142
Db 1027 TGTGGCTTACTTTACAGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAAGTGGCAGT 1086
QY 1143 GACCTTTCTGAGCGCTACTGGAAATTTCTCATTAGCAGTTCCGATCTTTTGTCCCTCT 1202
Db 1087 GACCTTTCTGAGCGGCTACTGGAAATTTCTCATTAGCAGTTCCGATCTTTTGTCCCTCT 1146
QY 1203 CTCTGCTTTGGCTTCCATGAACGGTGTGTGTTGCTGTCTCCAGGTTATTTCTATGTTGC 1262
Db 1147 CTCTGCTTTGGCTTCCATGAACGGTGTGTGTTGCTGTCTCCAGGTTATTTCTATGTTGC 1206
QY 1263 GTCTGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCGAAGCAGCACTCC 1322
Db 1207 GTCTGAGAGGGTCACTTCCAGAAATCTCTCCATGATTCATGTCGCGAAGCAGCACTCC 1266
QY 1323 TCTACAGCTGTTTATTTGTCACCTTTGACAATGATAATGCTCTCTCTGGAGACCT 1382
Db 1267 TCTACAGCTGTTTATTTGTCACCTTTGACAATGATAATGCTCTCTCTGGAGACCT 1326
QY 1383 CGACAGTCTTTTGAATTTTCTCAGTTTTCAGGTTGGCTTTTATTTATTTGGCTGGCAGTTGC 1442

TITLE Compositions and methods for the therapy and diagnosis of lung cancer

JOURNAL Patent: WO 0172295-A 440 04-OCT-2001;

CORIXA CORPORATION (US)

Location/Qualifiers

FEATURES 1..2239

Source /organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 619 a 444 c 493 g 683 t

ORIGIN

Query Match 100.0%; Score 2239; DB 6; Length 2239;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2239; Conservative 0; Mismatches 0; Gaps 0;

QY 1 GGAGGTGAAGTCAGCAGAGATCATGCCAGCCTGGGTGACAGTGCAGACTGCTCTCAAC 60
DB 1 GGAGGTGAAGTCAGCAGAGATCATGCCAGCCTGGGTGACAGTGCAGACTGCTCTCAAC 60
QY 61 AGAATTAAGGAAAAAGAAAGAAAGAGAGAGAGAGAAATCCAGGCCAAATGTGG 120
DB 61 AGAATTAAGGAAAAAGAAAGAAAGAGAGAGAGAGAAATCCAGGCCAAATGTGG 120
QY 121 CATAGATTTATCATATCTCGAATTTTGGATTTCTTTGGATTTCTTCTCATCATCTGATTC 180
DB 121 CATAGATTTATCATATCTCGAATTTTGGATTTCTTTGGATTTCTTCTCATCATCTGATTC 180
QY 181 GGAAGCCTGTGTGTCACCATCTCCAAAGAGAGTTACCTCAGGGAATGTTAACGG 240
DB 181 GGAAGCCTGTGTGTCACCATCTCCAAAGAGAGTTACCTCAGGGAATGTTAACGG 240
QY 241 AGGCTGCCCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCGCTTTTCAGGAAGAGAC 300
DB 241 AGGCTGCCCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCGCTTTTCAGGAAGAGAC 300
QY 301 GCCTTTTCAGGAAGAGACGCTTTTCAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGTC 360
DB 301 GCCTTTTCAGGAAGAGACGCTTTTCAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGTC 360
QY 361 ACTTTTACTAGGGAGTCTCCATATATTTGACCATCATTTGGAGCAGGAAATCTTCATC 420
DB 361 ACTTTTACTAGGGAGTCTCCATATATTTGACCATCATTTGGAGCAGGAAATCTTCATC 420
QY 421 TCTCCTAAGGGCGTCTCCAGAACACGGGAGCGGTGGGATGCTCTGACCATCTGGAGG 480
DB 421 TCTCCTAAGGGCGTCTCCAGAACACGGGAGCGGTGGGATGCTCTGACCATCTGGAGG 480
QY 481 GTGTGGGGTCTCTCAGTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATATA 540
DB 481 GTGTGGGGTCTCTCAGTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATATA 540
QY 541 AAGAAATCTGAGGTCATTAACATATATTTTGGAGTCTTTGGTCCATTTACCAAGCTTTT 600
DB 541 AAGAAATCTGAGGTCATTAACATATATTTTGGAGTCTTTGGTCCATTTACCAAGCTTTT 600
QY 601 GTACGAGTCTGGTGGAACTCTCATATACGCCCTGACGCTACTGCTGTATATCCCTG 660
DB 601 GTACGAGTCTGGTGGAACTCTCATATACGCCCTGACGCTACTGCTGTATATCCCTG 660
QY 661 GCATTTGGAGCCTACATCTTGGAAACCATTTTATTAATGTAATGTAATCTGAACTTGGG 720
DB 661 GCATTTGGAGCCTACATCTTGGAAACCATTTTATTAATGTAATGTAATCTGAACTTGGG 720
QY 721 ATCAAGCTCATTAACAGTCTGGCAATACATGATGCTGCTTAATAGCATGAGTGTGTC 780
DB 721 ATCAAGCTCATTAACAGTCTGGCAATACATGATGCTGCTTAATAGCATGAGTGTGTC 780
QY 781 AGCTGGAGCGCGGATTCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATA 840
DB 781 AGCTGGAGCGCGGATTCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATA 840
QY 841 ATTATAGTCCCTGGAGTTATGACAGCTAATTAAGGTCAACAGCAGAACCTTTAAAGACGCC 900
DB 841 ATTATAGTCCCTGGAGTTATGACAGCTAATTAAGGTCAACAGCAGAACCTTTAAAGACGCC 900

DB 841 ATTATAGTCCCTGGAGTTATGACAGCTAATTAAGGTCAACAGCAGAACCTTTAAAGACGCC 900
QY 901 TTTTTCAGGAAGAGATTCAGATTAATACCGCTTGCCACTGGCTTTTATTTATGTAATGTAT 960
DB 901 TTTTTCAGGAAGAGATTCAGATTAATACCGCTTGCCACTGGCTTTTATTTATGTAATGTAT 960
QY 961 GCATATGCTGCTGGTGTGTTTACCTCAACTTTGTTTACTGAAGAAGTAGAAAAACCCCTGAAAA 1020
DB 961 GCATATGCTGCTGGTGTGTTTACCTCAACTTTGTTTACTGAAGAAGTAGAAAAACCCCTGAAAA 1020
QY 1021 ACCATTTCCCTTGGCAATATGTAATCATCCATGCGCATTTGTCACCATTTGCTATGCTGACA 1080
DB 1021 ACCATTTCCCTTGGCAATATGTAATCATCCATGCGCATTTGTCACCATTTGCTATGCTGACA 1080
QY 1081 AATGTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCA 1140
DB 1081 AATGTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGGCA 1140
QY 1141 GTGACCTTTTCTGAGCGCTACTGGGAAATTTCTCATTTAGCAGTTCGATCTTTGTTGGC 1200
DB 1141 GTGACCTTTTCTGAGCGCTACTGGGAAATTTCTCATTTAGCAGTTCGATCTTTGTTGGC 1200
QY 1201 CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTCTCCAGGTTATTTCTATGTT 1260
DB 1201 CTCTCCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTCTCCAGGTTATTTCTATGTT 1260
QY 1261 CGCTCTCGAGAGGTCACCTTTCCAGAAATCCTCTCCATGATTCATGTCGCCAGCACACT 1320
DB 1261 CGCTCTCGAGAGGTCACCTTTCCAGAAATCCTCTCCATGATTCATGTCGCCAGCACACT 1320
QY 1321 CCTCTACAGCTGTTATTTGTTTTCACCCCTTTTGACAAATGATATGCTCTCTCTGGAGAC 1380
DB 1321 CCTCTACAGCTGTTATTTGTTTTCACCCCTTTTGACAAATGATATGCTCTCTCTGGAGAC 1380
QY 1381 CTCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTTGGCTTTTATTTGGGCTGGCAGTT 1440
DB 1381 CTCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTTGGCTTTTATTTGGGCTGGCAGTT 1440
QY 1441 GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCTGCTCTTCAAGGTGCCA 1500
DB 1441 GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCTGCTCTTCAAGGTGCCA 1500
QY 1501 CTGTTTCATCCAGCTTTGTTTTCCTTCACATGCTCTCTCATGTTGCTTCCCTCTAT 1560
DB 1501 CTGTTTCATCCAGCTTTGTTTTCCTTCACATGCTCTCTCATGTTGCTTCCCTCTAT 1560
QY 1561 TCGGACCCATTTAGTACAGGATTTGGCTTCGTCATCTGACTGAGTCCCTGCGCTAT 1620
DB 1561 TCGGACCCATTTAGTACAGGATTTGGCTTCGTCATCTGACTGAGTCCCTGCGCTAT 1620
QY 1621 TATCTCTTTTATATATGGCAAGAACCCAGGTGGTTTGAATAATATGTCAGAGAAATA 1680
DB 1621 TATCTCTTTTATATATGGCAAGAACCCAGGTGGTTTGAATAATATGTCAGAGAAATA 1680
QY 1681 ACCAGAAATTTACAATAATGCTGGAAGTTGTACCAAGAGAGATTAAGTTATGAACTAAT 1740
DB 1681 ACCAGAAATTTACAATAATGCTGGAAGTTGTACCAAGAGAGATTAAGTTATGAACTAAT 1740
QY 1741 GGACTTGGAGTCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTTTACTTCAT 1800
DB 1741 GGACTTGGAGTCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTTTACTTCAT 1800
QY 1801 TTTCTGAAAGTCTAGAGAAATACAACTTTTGGTGAATAACAAAAGAGTACGTTATTTT 1860
DB 1801 TTTCTGAAAGTCTAGAGAAATACAACTTTTGGTGAATAACAAAAGAGTACGTTATTTT 1860
QY 1861 TTCATATATTTTACGATATTCGAACTAATTTCTAAGAAATTTAGTTATACTCTATGTA 1920
DB 1861 TTCATATATTTTACGATATTCGAACTAATTTCTAAGAAATTTAGTTATACTCTATGTA 1920
QY 1921 TTATAGAAAGTCAATATGCACTTATTCATGAGTCCGCAATTTCTGAGTCTCTGATACC 1980
DB 1921 TTATAGAAAGTCAATATGCACTTATTCATGAGTCCGCAATTTCTGAGTCTCTGATACC 1980

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 03:54:58 ; Search time 6493 Seconds
(without alignments)
10035.605 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 2239

Sequence: 1 ggaggttgaaatgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2239	100.0	2239	6	AX321909	Sequence
2	2080.8	92.9	5981	6	AX321910	Sequence
3	1900.2	84.9	3144	9	HSX277882	AX321910 Homo sapi
4	1888.2	84.3	2155	9	AF212087	AX321910 Homo sapi
5	1888.2	84.3	2482	9	AF252872	AX321910 Homo sapi
6	1542.4	68.9	1874	9	AF200708	AX321910 Homo sapi
7	1540.8	68.8	1861	9	AB026891	AB026891 Homo sapi
8	1373.4	61.3	2000	9	AB040875	AB040875 Homo sapi
9	1208.6	54.0	2216	10	AB022345	AB022345 Mus muscu
10	529.2	23.6	64196	9	AC110804	AC110804 Homo sapi
11	379.2	16.9	3448	9	BC028216	BC028216 Homo sapi
12	379.2	16.9	6296	9	D87432	D87432 Human mRNA
13	367	16.4	2263	5	AF019906	AF019906 Xenopus l
14	366.4	16.4	1536	9	AB011263	AB011263 Homo sapi
15	366.4	16.4	2072	6	AR084930	AR084930 Sequence
16	366.4	16.4	2186	9	HSX018474	Y18474 Homo sapien
17	366.4	16.4	2237	9	HSX130718	Y130718 Homo sapi
18	366.4	16.4	2245	9	AF092032	AF092032 Homo sapi
19	365.4	16.3	1609	9	AF077866	AF077866 Homo sapi
20	365.4	16.3	1621	9	AB017908	AB017908 Homo sapi
21	364.8	16.3	2098	9	BC003062	BC003062 Homo sapi
22	364.8	16.3	2105	9	BC010107	BC010107 Homo sapi
23	364.8	16.3	2447	9	AB020532	AB020532 Homo sapi
24	362.2	16.2	1524	6	E41390	E41390 Amino acid
25	362.2	16.2	1767	9	AB018542	AB018542 Homo sapi
26	362.2	16.2	2254	5	XLASUR4B	Y12716 Xenopus lae
27	362.2	16.2	4539	6	E41369	E41369 Amino acid
28	362.2	16.2	4539	9	AB018009	AB018009 Homo sapi
29	362.2	16.2	4670	9	AF104032	AF104032 Homo sapi
30	357.8	16.0	1553	10	AF200684	AF200684 Rattus no
31	357.8	16.0	2077	10	AB020520	AB020520 Rattus no
32	354.4	15.8	2045	9	AK025377	AK025377 Homo sapi
33	351.4	15.7	4039	4	AF174615	AF174615 Bos tauru
34	349	15.6	2144	10	MMU0130943	AJ130943 Mus muscu
35	347.4	15.5	2168	10	BC014709	BC014709 Mus muscu
36	346.8	15.5	3423	10	AB023409	AB023409 Mus muscu
37	346.8	15.5	3441	10	AB017189	AB017189 Mus muscu
38	345.8	15.4	2158	10	MMU012754	AJ012754 Mus muscu
39	344.8	15.4	4117	6	E58469	E58469 Neutral ami
40	344.8	15.4	4117	10	AB024400	AB024400 Rattus no
41	343.8	15.4	3455	6	E31557	E31557 Neutral ami
42	343.8	15.4	3455	6	E41370	E41370 Amino acid
43	343.8	15.4	3455	10	AB015432	AB015432 Rattus no
44	343.8	15.4	3728	9	AF171669	AF171669 Homo sapi
45	343.8	15.4	4211	6	E58471	E58471 Neutral ami

ALIGNMENTS

RESULT 1	AX321909	AX321909	2239 bp	DNA	linear	PAT 15-DEC-2001
LOCUS	AX321909	Sequence 440 from Patent WO0172295.				
DEFINITION	AX321909	Sequence 440 from Patent WO0172295.				
ACCESSION	AX321909	AX321909				
VERSION	AX321909.1	GI:17906515				
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,					
	Indrias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,					
	Mannion,J. and Kalos,M.D.					

QY	340	GTGACGCTGAAGAGGAAGTCACTTTACTGAGGGAGTCTCCATTAATCAATTTGGCACCAC	399
Db	196	GTACCCCTGCAGCAATATCACACTGATCAATGGTGGCCATCATAGTGGCACCAC	255
QY	400	ATTGGAGCAGAAATCTCACTCTCCCTAAGGGGTCTCCAGAACACAGGGCAGGGTGGG	459
Db	256	ATCGGTTGGGCACTCTCGTGACCCCAAGGGGTCTCAAGGAAGCCGCTGGCCCGGA	315
QY	460	ATGCTCTGACCATCTGGACGGTGTGGGTCTGCTCACTATTGAGGCTTTGCTCTAT	519
Db	316	CTGTGCTGTGGGTGTGGGTGTGGGTCTTCTCCATGTGGGCGCACTGTGCTAC	375
QY	520	GCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCAATTAACATATATTTTGAAGTC	579
Db	376	CGGAGGTGGCACTACCATCTCAAGTCAAGGGCGCACTATGCCATACATGCTAGAGGTC	435
QY	580	TTTGGTCCATTAACAGCTTTTGTACGAGTCTGGGTGGAACCTCCTCATATACGCCCTGCA	639
Db	436	TACGGCTGCTGGCCGCTTCCCTCAAGCTCTGGATCGAGCTGCTCATTTCCGCCCTCC	495
QY	640	GCTACTGCTGTGATATCCCTGGCATTTTGGAGCTTACATTTCTGGAACCATTTTATTCAA	699
Db	496	TCACAGTACATCGTGGCGCTGGTCTTCGCCACATACCTGCTCAAGCGGCTTCCCGACT	555
QY	700	TGTGAATTCCTGAACCTTGGATCAAGCTCAATACAGCTGTGGGCATAACTGTAGTGATG	759
Db	556	TGTCCTGTCGCCGAGGAGGCTGCCAAGCTGTGGCTGCTCTCGCTGCTACTACTCAAG	615
QY	760	GTCTCAATAGCATGTGTCAGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGC	819
Db	616	GCTGTGAACCTGTACAGTGTGAAGGCTGTACCCGCTGTGAGGATGCTTTGGCGGTGCC	675
QY	820	AAGTCTACAGCAATTTCTGATAATATATAGTCCCTGGAGTTATGCAGCTAATTAAGGTCAA	879
Db	676	AAACTGTGCGCCCTGCGCCCTCATCATCTGCTCGGCTTCCATCCAGATGGGAAGGACATA	735
QY	880	-----ACGCAGAACTTTAAAGACCCCTTTTCAGGAGAGATTTCAAGT	921
Db	736	GGACAAGGGATGATCCAACTGACAGCAAGTTGCTCTTTGAAGGCACCAATCTGGAC	795
QY	922	ATTACGGGTGGCCATGCTGCTTTTATATGGAATGTATGCATATGCTGGCTGTTTAC	981
Db	796	GTGGGAACATTTGTTGGCATTTGTACAGTGGCTCTTCCCTACGAGGATGGAACAT	855
QY	982	CTCACTTTTGTACTGAAGAGTAGAAACCCCTGAAACCAATCCCTTTGCAATATGT	1041
Db	856	CTGAATTTTGTACGGAGGAGATGATCAACCCCTACAGGAACCTCCCTGGCCATCATC	915
QY	1042	ATATCCATGGCCATTTGTCACCATTTGCTATGCTGACAAATGTGGCTTACTTACGACC	1101
Db	916	ATCTCTTGGCCATTTGTCACCCCTGGTCTATGCTGACGAACCTGGCTTACTTCACTACC	975
QY	1102	ATTAATGCTGAGGAGTGTGCTTTTCAATGCACTGGCAGTACCTTTTCTGAGCGGCTA	1161
Db	976	CTGTCTACCAACCAAGATGCTGACATCTGAAGCGGTGGCTGTGATTTTGGAACTACCAC	1035
QY	1162	CTGGGAATTTCTCATTAGCAGTTCCGATCTTTGTTGGCTCTCCCTGCTTTGGCTCCATG	1221
Db	1036	CTGGAGTCAATGCTTCCGATCATCTCTGCTGCTGGGTGCTGCTGCTTCCGCTCTGTC	1095
QY	1222	AAGGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1281
Db	1096	ATGGGTCTCTGTTACAGTCTCAAGACTGTTCTGCTGGATCCAGGGGGGCCACCTG	1155
QY	1282	CCAGAAATCTCTCCATGATGATGTCGCCAGCAGCACTCTCTACAGCTGTATTGTT	1341
Db	1156	CCTTCCATCTCTCCATGATGATGTCGCCAGCAGCACTCTCTACAGCTGTATTGTT	1215
QY	1342	TTGCACCTTTGACATGATGATGCTCTCTGAGAGCTCGACAGCTTTTGAATTC	1401
Db	1216	ACGTGTGTCATGACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1275
QY	1402	CTCAGTTTTGCCAGGTGGCTTTTATTTGCGCTGGCAGTTGCTGGGCTGATTATCTCGA	1461

Search completed: April 16, 2003, 06:28:27
Job time : 533 secs

Db	1276	TTCAAGCTTCTCAACTGGCTGTGTGGCCCTGGCCATCATCGCATGATGTGGCTCGC	1335
QY	1462	TACAAATGCCAGATATGATCGTCTTTCAAGGTGCCACTGTTCATCCAGCTTTGTTT	1521
Db	1336	TTTAAGAAGCCTGAGCTGGAGGCTCCCATCAAGGTGAATCTGGCCCTCCAGTGTCTTT	1395
QY	1522	TCTTTCACATGCTCTTTCATGGTGGCTTTCCCTTATTCGGACCCCATTTTGTACAGG	1581
Db	1396	ATCCTGGCTGGCTCTTCTCTCATCGCCGTGCTCTTCTGGAAGACACCCCTGGAGTGGC	1455
QY	1582	ATTGGCTTGGTCAATCACTCTGACTGGAGTCCCTGGGTATTTATCTCTTTATATATGGAC	1641
Db	1456	ATTGGCTTGGCCATCATCTCAGCGGCTGCTCTACTTCTTTGGTGTGTGGTGGAA	1515
QY	1642	AAGAAACCCAGGTGGTT	1658
Db	1516	AACAAGCCCAATGGAT	1532

XX	Sequence	4585 BP; 794 A; 1475 C; 1303 G; 1013 T; 0 other;	
QQ	Query Match	15.8%; Score 353.6; DB 22; Length 4585;	
QQ	Best Local Similarity	55.2%; Pred. No. 1.6e-82;	
QQ	Matches	736; Conservative 0; Mismatches 589; Indels 9; Gaps 2;	
QY	332	GAGAGAAAGTGCAGCTGAAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATG 391	
DB	187	GCAGGGCGTACCTCGAGGGGAACATCACGCTACTCAAGGGCGTGGCGTTCATGTTG 246	
QY	392	GCACCATATTTGGAGCAGGAACTTTCATCTCTCAAGGCGTGGCTCCCAAGACCGGCA 451	
DB	247	TCGCCATATGGGCTCGGGCATCTTCTGAGCCGACCGGGCGTGTCAAGAGGAGGAGCT 306	
QY	452	GCCTGGGCATCTCTCTGACCATCTGGACGCTGTGGGGTCTCTGCTCACTATTTGGAGCTT 511	
DB	307	CGCCGGGGCTGGCGCTGGTGTGGCGCGCTGCGGCGTCTTCCATCTGCTGGCGCGC 366	
QY	512	TGCTTTATGCTGAATTTGGGAACAACTATAAGAAATCTGGAGTCTATTACACATATTT 571	
DB	367	TCTCTACGGGGAGCTGGGACCAACCATCTCCAAATCGGGCGGCACTACGCCCTACATGC 426	
QY	572	TGGAAGTCTTTGGTCCATACAGCTTTTGTACAGTCTGGGTGGTGAACCTCTCATATAC 631	
DB	427	TGGAGTCTACGGCTGCTGCCCGCTTCTCAAGCTCTGGATCGAGCTGCTCATCATCC 486	
QY	632	GCCCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATTTCTGGAACCATTT 691	
DB	487	GGCTTTCATCGAGTACATCGTGGCCCTGCTCTTCCGCCACTTACCTGCTCAAGCCCTCT 546	
QY	692	TTATTTCAATGTAAATCCCTGAACTTTCGATCAAGTCTATTACAGTGTGGGCATAACTG 751	
DB	547	TCCCCACCTCCCGGTCGCCGAGGAGCAGCAAGCTCGTGGCTGCTGCTGCTGCTGCTG 606	
QY	752	TAGTATGGTCTTAATAGATGAGTGTGAGTGTGGAGCGCCGCGATCCAGATTTCTTAA 811	
DB	607	TGCTACGGCGTGAATGCTACAGGTGAAGCGCGCCACCGGGTCCAGGATGCGCTTTG 666	
QY	812	CTTTTTCAGAGCTCACAGCAATTTCTGATTAATATAGTCTCCCTGGAGTATGCACTAATTA 871	
DB	667	CCGCGGCAAGCTCTCGCCCTGGCTGATCATCTGCTGGGCTTCTGCTCCAGATCGGA 726	
QY	872	AAGTCAACGCGAGAACTTTAAGAGCGCTTTTACAGAGAGATTTCAAG-----TATTA 925	
DB	727	AGGTGATGTGTCCAACTCTAGATCCCAACTTCTCATTTGAAGGCAACCAACTGATGG 786	
QY	926	CGCGTTGCCACTGGCTTTTATTATGGAATGTATCATATGCTGGCTGCTGTTTACCTCA 985	
DB	787	GGACATTTGCTGGCATATACAGGGCTCTTTGCCCTATGAGGATGGAATTACTTGA 846	
QY	986	ACTTTGTTACTGAAGAGTGAAGAAACCTGAAAAACCATTTCCCTTGGCAATATGTATAT 1045	
DB	847	ATTTCGTCACAGAGAAATGATCAACCCCTACAGAAACCTGCCCCCTGGCCATCATCATCT 906	
QY	1046	CCATGGCCATTTACCATTTGGCTATGTCTGACAAATGTGGCTTACTTTAGACCAATTA 1105	
DB	907	CCCTGGCCATCTGGAGCTGGTGTAGTGTGCTGACCAACCTGGCTTACTTACCAACCTGT 966	
QY	1106	ATGCTGAGGAGCTGCTGCTTTCAATGTCAGTGGCTGACCTTTTCTGAGCGGCTACTGG 1165	
DB	967	CCACCGAGAGATGTGCTGCTCGAGGCGTGGCCGTGGACTTCGGGAATATACCTGTG 1026	
QY	1166	GAAATTTCTCATTAGCAGTTTCGATTTTGTGGCCCTCTCTGCTTGGCTCCATGAAGC 1225	
DB	1027	GCCTAATGTCTGATCATCCCGTCTCTGCTGGCCCTGCTGCTGCTGCTCGCTCAATG 1086	
QY	1226	GTGGTGTGTTGCTGCTCAGGTATTTCTATGTTCCGCTCTGAGAGGCTGACCTTCCAG 1285	
DB	1087	GGTCCCTGTTTCACTCATCTCCAGGCTCTTCTGTTGGGGTCCCGGGAAGGCCACCTGCCCT 1146	
QY	1286	AAATCTCTCTCATGATTTATGTCCGCAAGCACACTCTCTTACAGCTGTTATTGTTTTC 1345	
DB			
DB	1147	CCATCCTCTCCTCATGATCCACCCACAGCTCTCTACCCCGTGGCTGCTGCTGTTTCACT 1206	
QY	1346	ACCCTTTGCAATGATGATGCTTCTCTGAGAGCTCGACAGTCTTTTGAATTTCCCTCA 1405	
DB	1207	GTGTGATGACGCTGTTTACGGCTTCTCAAGGACATCTTCTCCGTCATCAACTTCTTCA 1266	
QY	1406	GTTTGGCCAGGTGGCTTTTATTGGGCTGGCAGTTGCTGGGCTGATTTATCTTTCGATACA 1465	
DB	1267	GCTTCTCAACTGGCTTCTGCTGGCCCTGGCCATCATCGGCATGATCTGGCTGGCCACA 1326	
QY	1466	AATGCCAGATATGATGCTCTTCAAGGTGCCACTGTTTATCCAGCTTGTGTTTCTT 1525	
DB	1327	GAAAGCTGAGCTTGAGCGGCCCATCAAGGTGAACCTGGCCCTGCTGTGTTCTTTCATCC 1386	
QY	1526	TCACATGCTCTCTCATGTTGGCTTTCCTCTCT---ATTCCGAGCCATTTAGTACAGGA 1582	
DB	1387	TGGCTGCTCTCTCTGATCGCGTCTCTTCTTGAAGACACCCCGTGGAGTGTGGCAT 1446	
QY	1583	TTGGCTTCTCTCATCTCTGACTGGAGTCCCTGCGTATTTATCTTATATATATGGGACA 1642	
DB	1447	CGGACTTCCACCATCATCTCTCAGCGGCTGCCGCTCTACTTCTTCCGGGCTGCTGGTGGAAAA 1506	
QY	1643	AGAAACCCAGGTGG 1656	
DB	1507	ACAAGCCCAAGTGG 1520	
XX	RESULT 14		
XX	AAAF77276		
ID	AAAF77276	standard; cDNA; 4117 BP.	
XX	AAAF77276;		
XX	24-MAY-2001	(first entry)	
DE	Rat	cDNA encoding L-type amino acid transporter 2 (LAT2).	
KW	Neutral	amino acid transporter; amino acid transport activating factor;	
KW	4F2hc;	4F2 heavy chain; drug permeation; poison; rat; LAT2;	
XX	L-type	amino acid transporter 2; ss.	
OS	Rattus	sp.	
XX	JP2000342270-A.		
XX	12-DEC-2000.		
XX	04-JUN-1999;	99JP-0158571.	
XX	04-JUN-1999;	99JP-0158571.	
XX	(KAGA-)	KAGAKU GIJUTSU SHINKO JIGYODAN.	
XX	WPI;	2001-260288/27.	
XX	P-PSDB;	AAB72393.	
PT	A	neutral amino acid transporter having broad substrate selectivity	
PT	useful	for in vitro analysis of permeation of drugs and poisons through	
PT	cell	membranes, blood and tissues	
XX	Claim	7; Page 12-17; 37pp; Japanese.	
XX	This	invention relates to a neutral amino acid transporter having broad	
XX	substrate	selectivity to neutral amino acids, and which is inactivated by	
XX	an	amino acid transport activating factor 4F2hc (4F2 heavy chain). The	
XX	neutral	amino acid transporter and its gene are useful in vitro	
XX	analysis	of permeation of drugs and poisons through cell membrane, blood	
XX	and	tissues. The present sequence represents cDNA encoding the rat L-type	
XX	amino	acid transporter 2 (LAT2).	
XX	Sequence	4117 BP; 920 A; 1206 C; 970 G; 1021 T; 0 other;	
XX	Query Match	15.4%; Score 344.8; DB 22; Length 4117;	

Db 672 CAAATGGGAACCTGGTACAGATATTTTACCTATGCTAAAGTATTGGCAGCTGATCG 731
 QY 840 AATTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCAACGCGCAACTTTAAAGACGC 899
 Db 732 GGTCACTGTTGACAGGATTTAGACTTGGCCAGGAGCCTTACTCATTTTGAAGATTC 791
 QY 900 CTTTTCAGGAAGAGATCAAGTATTACGGGGTTGCCACCTGGCTTTTATTATGGAATGTA 959
 Db 792 CTTTGGGGTTTCATCTTTGCACTGGGTGACATTTGCCCTGGCACCTACTCAGCTCTGTT 851
 QY 960 TGCATATGCTGGCTGGTTTACCTCACTTGTGTTACTGAAGAGTAGAAAAACCTGAAAA 1019
 Db 852 CTCCTACTCAGGCTGGACACCTCACTATGTCTCACTGAAGAGATCAAGAACTCTGAGAG 911
 QY 1020 AACCATTTCCCTTGCATATGATATATCCATGGCCATTTGACCATTTGGCTATGCTGAC 1079
 Db 912 GAACCTGGCCCTCTCCATTTGGCATCTCCATGGCCCTTGTGACCATCATCTATCTTGAC 971
 QY 1080 AAATGTGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTCTTTCAATGCACTGGC 1139
 Db 972 CAATGTGGCTATATACCTGTGCTAGACATGAGAGACATCTTGGCCAGTGATGCTGTTGC 1031
 QY 1140 AGTGACCTTTTTCAGGGCTACTGGGAATTTCTCATTTAGCAGTTCCCATCTTTGTTGC 1199
 Db 1032 TGTGACTTTTTCAGAGATCAGATATTTTGAATATTTAACTGGATAATTTCCACTGTCTGAT 1091
 QY 1200 CCTCTCTCTGTTGGCTTCCATGAACGCTGGTGTGTTGCTGCTCCAGCTTATCTCATGT 1259
 Db 1092 ATTATCTGTTTGGTGGCTCAATGCCCTCAATGTTGGCTGCTTCTAGGCTTTCTTTGT 1151
 QY 1260 TGGCTCTCGAGAGGCTCACCTCCAGAAATCTCTCCATGATCATGTCGCGAAGCACAC 1319
 Db 1152 GGGCTCAAGAGAAGGCGCATCTCCCTGATGCCATCTGCATGATCCATGTTGAGCGTTTCA 1211
 QY 1320 TCCTCTACCAAGCTTATGTTTGTGACCTTTGACAAATGATGATGCTCTTCTCTGGAGA 1379
 Db 1212 ACCAGTGGCTTCTCTGCTCTCAATGATATCATGTCATGTCATGCTGCTGGGAAGA 1271
 QY 1380 CTTGACAGCTTTTGAATTTCTCAGTTTGGCCAGTGGCTTTTATTTGGGCTGGCAGT 1439
 Db 1272 CATCTTCCAGCTCATTAATCTACAGCTTACAGCTTACAGCTTCTTGTGGGCTTCTAT 1331
 QY 1440 TGTGGGCTGATTTATCTTCGATACAAATGCGCAGATATGATGCTCTTTTCAAGTGGC 1499
 Db 1332 TGTGGGTCAGCTTTATCTGGCTGGAAGAGGCTGATCGACCTCGTCCCTCAAGCTCAG 1391
 QY 1500 ACTGTTTCATCCAGCTTTGTTTCTTTCATGATGCTTCTTCTGCTGCTTCTTCCCTCTA 1559
 Db 1392 CGTTTCTTCCGATTTGCTTCTGCTCTGACCATCTTCTGCTGCTGCTGCTGCTGCTT 1451
 QY 1560 TTCGACCCATTTAGTACAGGATTTGGCTTCTGCTCATCTCTGCTGCTGCTGCTGCTGCT 1619
 Db 1452 CAGTGATACATCACTCCCTCATCGGATTTGCCATTTGCCCTCTCAGGCTGCTCTTTA 1511
 QY 1620 TTATCTCTTTTATATA 1635
 Db 1512 CTTCTCATCATCAGA 1527

RESULT 12
 AAA08393
 ID AAA08393 standard; RNA; 1524 BP.
 XX
 AC AAA08393;
 XX
 DT 12-JUL-2000 (first entry)
 XX
 DE Human L-type amino acid transporter 1 RNA sequence SEQ ID NO:26.
 XX
 KW L-type amino acid transporter 1; LAT1; 4F2hc; 4F2 heavy chain;
 KW neutral amino acid transporter; tumour; cancer; proliferation;
 KW cell membrane surface 4F2 molecule; anticancer; ss.

XX Homo sapiens.
 OS WO200014228-A1.
 PN 16-MAR-2000.
 PD 03-SEP-1999; 99WO-JP04789.
 PF 03-SEP-1998; 98JP-0249993.
 PR 02-SEP-1999; 99JP-0248546.
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 PI Endou H, Kanai Y;
 XX WPI: 2000-256979/22.
 DR P-PSDB: AAY82488.
 XX Neutral amino acid transporter protein which conjugates with cell membrane protein 4F2 and operates independently of sodium ions is useful for screening potential cancer proliferation inhibitors -
 Claim 13; Page 176-180; 189pp; Japanese.
 The present invention describes a cell surface protein which is an L-type amino acid transporter-1 (LAT1), which mediates the transport of neutral amino acids, leucine, isoleucine, phenylalanine, methionine, tyrosine, tryptophan, valine and histidine, into the cell independently of sodium ions. The LAT1 protein conjugates with the cell membrane surface molecule 4F2. Antisense DNA or RNA, antibodies, antibody fragments, and peptide and non-peptide antagonists to LAT1 are useful as anticancer agents. The present sequence encodes human LAT1, which is specifically claimed in the present invention.
 XX Sequence 1524 BP; 245 A; 516 C; 438 G; 325 U; 0 other;

Query Match 16.2%; Score 362.2; DB 21; Length 1524;
 Best Local Similarity 38.0%; Pred. No. 4.8e-85;
 Matches 506; Conservative 228; Mismatches 593; Indels 6; Gaps 1;
 QY 332 GAGAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTAGGGAGTCTCCATTATCATTTG 391
 Db 122 GCGAGGGGUGACCCUGGCGGGAACAUACGCGUGCUAACGCGGUGGCCAUCACUGG 181
 QY 392 GCACCATCATTTGGAGCAGGAATCTCTCTCTAAGGGCTGCTCCAGAACACGGGCA 451
 Db 182 GGACCAUUAUUGGCGGCAUCUUGAGCGCCACGCGGUGGUCUACAGGAGGAGCU 241
 QY 452 GCGTGGGATGCTCTGACCATCTGACGGTGTGTGGGGTCTCTGCTACTATTGAGGTT 511
 Db 242 CGCCGGGCGUGGCGUGGUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
 QY 512 TGTCTTATGCTGAATTTGGGAACAACATAAAGAAATCTGGAGGTCTTACACATATATT 571
 Db 302 UGCUACGCGGAGCUGCGGACCAACCAUCCAAUCCGCGGCGGCGGCGGCGGCGGCGG 361
 QY 572 TGGAAAGTCTTTGGTCCATTACCAAGCTTTTGTACGAGTCTGGTGGAACTCTCTAATAAC 631
 Db 362 UGGAGGUCUACGCGGUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 421
 QY 632 GGCCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGAGCTTACATTTCTGGAACCAATTT 691
 Db 422 GGCCTGCAGCTACTGCTGTGATATCCCTGGCATTTGGAGCTTACATTTCTGGAACCAATTT 691
 QY 692 TTATTCAATGTGAATCCCTGAACTTGGATGCAAGCTTATACAGCTTGGGCGGCGGCGG 751
 Db 482 UCCCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541
 QY 752 TAGTGATGCTCTTAATAGCATGAGTGTACGCTGAGGCGCGGCGGCGGCGGCGGCGGCGG 811
 Db 542 UGCUCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601

QY 900 CTTTTCAGGAAGAGATTCAGTATTACGGGTTGCCACTGGCTTTTATTATGTAATGTA 959
 Db 792 CTTTGGGGTTTCATCATTTGCACTGGTGGTACATTCGCCCTGGCACTGTACTCAGCTCTGTT 851
 QY 960 TGCATATGCTGGCTGGTTTACTCACTTGTACTGAAGAGTAGAAAACCCCTGAAAA 1019
 Db 852 CTCCTACTCAGCTGGGACACCTCACTATGTCTACTGAAGAGATCAAGAATCCTTGAGAG 911
 QY 1020 AACCATTCCTCTTGGCTCCATGATATATCCATGCCATTTGCACCAATGGCTATGTGCTGAC 1079
 Db 912 GAACTGCCCCCTCTCATTTGGCATCTCCATGCCCATTTGTCCACATCATATATCTTGAC 971
 QY 1080 AAATGTGGCTACTTTAGCAGCAATTAATGCTGAGGAGTGTGCTTTTCAAAATGCAAGTGGC 1139
 Db 972 CAATGTGGCTATTATCTGTCTAGACATGAGACATCTTGGCCAGTGTGCTGTGTC 1031
 QY 1140 AGTGACCTTTCTGAGCGGCTACTCGGAAATTTCTCATTTAGCAGTTCGGAATCTTTGTTGC 1199
 Db 1032 TGTGACTTTTGGAGATCAGATATTTTGAATATTTAACTGGGATTAATTCACACTGTGAGTGC 1091
 QY 1200 CTTCTGCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTTATCTATGT 1259
 Db 1092 ATATCTCTGTTTGGTGGCTCAATGCCCTCCATGTGGCTGCTCTAGGCTTTTCTTTGT 1151
 QY 1260 TCGCTCTCGAGAGGCTACCTTCCAGAAATCCTCTCCATGATTCATGTCCGCAAGCACAC 1319
 Db 1152 GGGCTCAAGAGAAGGCCATCTCCCTGATGCCATCTGCATGATCCATGTTGAGCGGTTAC 1211
 QY 1320 TCCTCTACAGCTGTTATTTGTTTCCACCCCTTGAACATGATAATGCTCTCTCGGAGA 1379
 Db 1212 ACCAGTGGCTTCTGCTCTTCAATGGTATCATGGCATTTGATCTACTTGTGCGTGAAGA 1271
 QY 1380 CTTGACAGCTCTTTGAATTTCTCAGTTTTCAGAGTGGCTTTTATTTGGGCTGGCAGT 1439
 Db 1272 CATCTCAGCTCAATTAATCTACAGCTTCAGCTTGGTCTTCTTGTGGGCTTTCTAT 1331
 QY 1440 TGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATGCTCTTTCAGAGGTGCC 1499
 Db 1332 TGTGGGCTCAGCTTTATCTGCGTGGAGAGGCCCTGATCGACCTCTGCCCTCAAGCTCAG 1391
 QY 1500 ACTGTTTATCCAGCTTTGTTTCTTCCATGCGCTTTCATGCTTTCATGTTGCGCTTCCCTCTA 1559
 Db 1392 CGTTTCTTCCCGATTTGCTTCTGCTCTGCACCATCTTCTGTTGGCTGTCTCCACTTTA 1451
 QY 1560 TTCGACCCATTTAGTACAGGATTTGGCTTCTGTCATCACTCTGACTGAGTCCCTGGCTA 1619
 Db 1452 CAGTATATCACTCCCTCATCGGCAATGGCAATGCCCTCTCAGGCGCTTCCCTTTTA 1511
 QY 1620 TTATCTCTTTTATA 1635
 Db 1512 CTTCTCATCATCAGA 1527

RESULT 11

AAX86999
 ID AAX86999 standard; cDNA; 2072 BP.

AC AAX86999;
 XX

DT 18-OCT-1999 (first entry)

XX Human amino acid permease homologue (AAPH) nucleotide sequence.

KW Amino acid permease homologue; AAPH; cancer; inflammatory disorder;
 KW autoimmune disorder; cell proliferation disorder; adenocarcinoma;
 KW AIDS; Crohn's disease; rheumatoid arthritis; arteriosclerosis; human;
 KW psoriasis; hepatitis; gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FH 142..1677

FT CDS /*tag= a

FT /product= "amino acid permease homologue (AAPH)"
 PN US942399-A.
 XX
 PD 24-AUG-1999.
 XX
 PP 06-MAY-1998; 98US-0073362.
 XX
 PR 06-MAY-1998; 98US-0073362.
 XX
 PA (INCY-) INCYTE PHARM INC.

Corley NC, Hillman JL, Yue H;

WPI; 1999-493508/41.

P-PSDB; AAY27076.

New human amino acid permease homologue (AAPH) polynucleotide and its complement useful for prevention and treatment of cancer, autoimmune and cell proliferation disorders

Claim 4; Fig 1A-F; 33pp; English.

This sequence encodes a human amino acid permease homologue (AAPH). The protein can be expressed by standard recombinant methodology. AAPH polynucleotides and polypeptides are useful for diagnosis, treatment and prevention of cancer e.g. adenocarcinoma, lung cancer, ovarian cancer; inflammatory/autoimmune disorders e.g. AIDS, Crohn's disease, rheumatoid arthritis; and cell proliferation disorders e.g. arteriosclerosis, psoriasis, hepatitis. AAPH polynucleotides are useful in gene therapy, in PCR-based methods for detecting upstream sequences, and also for generating hybridisation probes for mapping the naturally occurring genomic sequence.

Sequence 2072 BP; 446 A; 537 C; 494 G; 595 T; 0 other;

Query Match 16.4%; Score 366.4; DB 20; Length 2072;
 Best Local Similarity 54.6%; Pred. No. 4.4e-86;
 Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;

QY 300 CGCCTTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAAAGTSCAGCTGAAGAGGAAAGT 359

Db 192 CTCCTCTTGGTGTATGGGCGCAGCCAGCGCGGAGCAGGTGAAGCTGAAGAAGGAGAT 251

QY 360 CACTTTACTCAGGGGAGTCTCCATTTATCATTTGGCACCATCATTTGAGCAGGATCTTCAT 419

Db 252 CTCACCTGCTTAACGGGCTGTGCTGATTTGGGGAACATCATGGGCTCAGGCACTTTTGT 311

QY 420 CTCCTCTAAGGCGGTGCTCCAGAACACAGCGGCGGCGATGCTCTGACCATCTGGAC 479

Db 312 TTCCCCCAAGGTTGTCTCATATACAGTGCCTCTTTGGTCTCTCTCTGTCATCTGGGC 371

QY 480 GGTGTGGGGTCTCTCACTATTTCAGGCTTTGTCTTATGCTGAATTTGGGAACAATAT 539

Db 372 TGTGCGGGGCTCTCTCCGCTCTTTGGGCGCTTTCTTATGCGGAACCTGGCACCACCAT 431

QY 540 AAGAAATCTGGAGGTCATTACACATATATTTTGGGAAGTCTTTGGTCCATTACCACTTTT 599

Db 432 TAAGAAATCTGGGCGCAGCTATGCTTATCTGAGGCGCTTTGGAGGATTCCTTGTCTTT 491

QY 600 TGTCAGGCTCTGGGTGGAACTCTCATATACGCCCTCGAGCTACTGCTGTGATATCCCT 659

Db 492 CATCAGACTCTGACCTCTCCCTCATCATTTGAGCCACCCAGCCAGGCCATCATTTGCCAT 551

QY 660 GGCAATTTGGACGCTACATCTTGGAAACCATTTTATTCATATGGAATCCCTGAACTTGC 719

Db 552 CACCTTTGGCAACTACATGTCACAGCTCTCTTCCCGAGCTGCTTGGCCCTTATGCTGTC 611

QY 720 GATCAAGCTCATACAGCTCTGGGCATACCTAGTAGTATGGTCTCTAAATAGCATGAGTGT 779

Db 612 CAGCCCGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671

QY 780 CAGCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTTCGAAGCTCACAGCAATTCGTAT 839

Db 816 GAACCTGCCCCCTCTCCATTGGCATCCCATGCCATTGTGACCATCATATATCTTGAC 875
QY 1080 AATGTGGGCTACTTTACACACATTAATGCTGAGAGCTGCTGCTTTCAATGCAAGTGGC 1139
Db 876 CAATGTGGGCTATTATACCTGCTAGACATGAGACACATCTTGGCCAGTGATGCTGTTGC 935
QY 1140 AGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTCCGATCTTTGTTGC 1199
Db 936 TGTGACTTTTTCAGATCAGATATTTTGAATATTTAACTGGATATTTCCACTGTCAGTTC 995
QY 1200 CTTCTCTCTGCTTTGGCTCCATGAAGGGTGGTGTGTTGCTGCTCCAGGTTATCTATGT 1259
Db 996 ATTATCTCTTTTGGTGGGCTCAATGCCCTCCCATTTGGCTGCTCTAGGCTTTTCTTGT 1055
QY 1260 TGGCTTCAGAGGGGTCACCTTCCAGAAATCCTCTCCATGATTCATGTCGCCGACACAC 1319
Db 1056 GGGCTCAAGAGAGGGCCATCTCCCTGATGCCATCTGCTATGATTCATGAGCGGTTCCAC 1115
QY 1320 TCCTCTACAGGCTGTTATTTTGGCACCTTTTGACAATGATGATGCTCTCTCTGGAGA 1379
Db 1116 ACCAGTGCCTTCTCTGCTCTTCAATGGTATCATGCAATGATCTACTTGTGGTGAAGA 1175
QY 1380 CTTGACAGCTCTTTGAATTTCTCAGTTTTCGAGGTGGCTTTTATTTGGCTGGCAGT 1439
Db 1176 CATCTTCCAGCTCAATTAACCTACAGCTTCAGCTTCTGTTCTTTGTTGGGCTTTCTAT 1235
QY 1440 TGGTGGGCTGATTATCTTCGATACAAATGCCAGATATGCTGCTCTTTTCAAGTGGC 1499
Db 1236 TGTGGGTGAGCTTATCTGCGTGGNAGGAGCTGATCGACCTCTGCTCCCTCAAGCTAG 1295
QY 1500 ACTGTTTATCCAGCTTTGTTTCTCTTCACATGCTCTTCATGTTGGCTTTCCCTCTA 1559
Db 1296 CGTTTCTTCCGAGTGTCTCTGCTCTGACCATCTCTCTGCTGCTGCTGCTGCTTCA 1355
QY 1560 TTCGGACCATTTAGTACAGGATTTGGCTTCTGCTATCATCTGACTGAGTCCCTGCGTA 1619
Db 1356 CAGTGATCTATCACTCCCTCATGGCATTTGCCATTTGCCCTCTCAGGCTTCCCTTTTA 1415
QY 1620 TTATCTCTTATATATA 1635
Db 1416 CTCTCTCATCATCAGA 1431

RESULT 10

ID AAZ23240
AAZ23240 standard; cDNA; 2072 BP.

XX AAZ23240;

AC AAZ23240;

XX 24-JAN-2000 (first entry)

XX Human amino acid permease homolog (AAPH) nucleic acid sequence.

XX Amino acid permease homolog; AAPH; cancer; inflammatory disorder; AIDS;
KW autoimmune disorder; anemia; asthma; atherosclerosis; diabetes mellitus;
KW multiple sclerosis; osteoarthritis; osteoporosis; rheumatoid arthritis;
KW ulcerative colitis; infection; cell proliferation disorder; human; ss.

XX Homo sapiens.

XX US981242-A.

PN 09-NOV-1999.

XX 03-FEB-1999; 99US-0243920.

XX 06-MAY-1998; 98US-0073362.

XX (INCY-) INCYTE PHARM INC.

XX Yue H, Corley NC, Hillman JL;

XX WPI; 1999-63325/54.

DR P-PSDB; AAY41159.

XX New human amino acid permease homolog, useful in the diagnosis,
PT treatment and prevention of cancer, inflammatory/autoimmune disorders
XX and cell proliferation disorders -

XX Disclosure; Fig 1A-F; 32pp; English.

XX This cDNA encodes a human amino acid permease homolog (AAPH). The AAPH
CC nucleic acid, protein, AAPH (antagonists are useful in the diagnosis,
CC treatment and prevention of cancer such as adenocarcinoma, leukemia,
CC lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and cancers of the
CC adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall
CC bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung,
CC muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands,
CC skin, spleen, testis, thymus, thyroid or uterus; inflammatory/autoimmune
CC disorders such as AIDS, Addison's disease, adult respiratory distress
CC syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia,
CC asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune
CC thyroiditis, bronchitis, cholecystitis, contact dermatitis, diabetes
CC mellitus, Crohn's disease, atopic dermatitis, dermatomyositis,
CC emphysema, multiple sclerosis, osteoarthritis, osteoporosis, rheumatoid
CC arthritis, ulcerative colitis, viral, bacterial, fungal, parasitic,
CC protozoal infections; and cell proliferation disorders such as acinic,
CC keratosis, cirrhosis, arteriosclerosis, myelofibrosis, and primary
CC thrombocythemia.

XX Sequence 2072 BP; 446 A; 537 C; 494 G; 595 T; 0 other;

XX Query Match 16.4%; Score 366.4; DB 20; Length 2072;
Best Local Similarity 54.6%; Pred. No. 4.4e-86;

XX Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;

QY 300 CGCCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAGAGAGAGAGAGAGT 359
Db 192 CTCCTTTTGGGTGATGGGCGCCAGCCAGGCGGAGAGGTGAAGCTGAAGAGAGAGAT 251

QY 360 CACTTTTACTAGGAGGTCTCCATTATTTGGCCACCATCATTTGGAGCAAGAAATCTTCAT 419
Db 252 CTCCTGCTTAACGGCGTGTGCTGATTTGGGGAACATGATCGGCTCAGGCATCTTGT 311

QY 420 CTCCTTAAGGGGTGCTCCAGAACACAGGCGGAGGTGGCATGCTCTGACCATCTGAC 479
Db 312 TTCCTCCCAAGGGGTGCTCATATACAGTGCCTCTCTTTGGTCTCTCTGCTGCTGGC 371

QY 480 GGTGTGGGCTCTCTGCTCACTATTTGGAGCTTTGCTTATGCTGAATTTGGGAACAATAT 539
Db 372 TGTGGGGGCTCTCTCTGCTCTTTGGGGCCCTTTGTTATGCGGAAGTGGCACCACAT 431

QY 540 AAGAAATCTGGAGGTCAATTACACATATATTTTGAAGTCTTTTGGTCCATTTACAGCTTT 599
Db 432 TAAGAAATCTGGGCGCAGCTATGCTATATCTGAGGCGCTTTGGAGGATTCCTTGTCTT 491

QY 600 TGTAGAGTCTGGTGGAGTCTCCATATAGCCCTGACCTGCTGCTGATATCCCT 659
Db 492 CATCAGACTGGAGCTCCCTGCTCATCATTTAGCCCAACAGCAGGCGCATATGGCAT 551

QY 660 GGCATTTGGAGCGCTACATTTCTGGAACCACTTTTATTAATGTAATCCCTGAACCTGC 719
Db 552 CACCTTTGCCAACTACATGTTACAGCTCTCTTCCCGAGCTGCTTCGCCCTTATGCTGC 611

QY 720 GATCAAGTCAATTACAGCTGTGGGCAATACTGTAGTGTGCTTCTTAATAGCATGAGTGT 779
Db 612 CAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671

QY 780 CAGCTGGAGCGCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATCTGAT 839
Db 672 CAAATGGGGAACCTGTTGACAGATATTTTCCCTATGCTAAAGTATTTGGCAGCTGATCG 731

QY 840 AATTATAGTCCCTGGAGTTATGACAGTAAATTAAGGTCAACACGACAGCAATTTAAGACGC 899
Db 732 GGTATGCTTGCAGCATTTGTAGACTTGGCCAGGAGCCCTTACTCATTTTGAATTC 791

Db 831 CAATGGCCCTATTATCTGTCTAGACATGAGACATCTTGGCCAGTGATGCTGTGC 890
 Qy 1140 ACTGACCTTTCTGAGCGCTACTGGAAATTCATACAGCTCCGATCTTGTTC 1199
 Db 891 TGTGACTTTTGCAGATAGATTTGGAAATATTTAACTGGATAATTCACATGTCAGTTC 950
 Qy 1200 CCTCTCCCTGCTTTGGCTCCATGAACGGTGGTGTCTCTCTCCAGGTTATCTATGT 1259
 Db 951 ATTATCCCTGTTTGTGGCCCAATGCCCTCCATTTGGCTGCTCTAGGCTTTCTTGT 1010
 Qy 1260 TCGGCTCTGAGAGGTACCTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGACAC 1319
 Db 1011 GGGCTCAAGAGAAGGCCATCTCCCTGATGCCATCTGCAATGATGTCGCGGTTTCA 1070
 Qy 1320 TCCTCTACCAAGCTGTTATTTGTTTTCACCCCTTGACAATGATAATGCTCTCTCTGGAGA 1379
 Db 1071 ACCAGTGGCTTCTGTCTCTCAATGGTATCATGGCATTGATCTGTGCGTGAAGA 1130
 Qy 1380 CCTGACAGCTCTTTTGAATTCCTCAGTTTTCAGGTTGGCTTTTATTTGGCTGGCAGT 1439
 Db 1131 CATCTTCCAGCTCATTAATACTACTACAGCTTCAGCTACTGGTCTTTTGGGGCTTTCTAT 1190
 Qy 1440 TCGTGGCTGATTTATCTTCGATACAAATGCGCAGATGATGATCGTCTCTTCAAGGTGCC 1499
 Db 1191 TGTGGTCAAGCTTTATCTGCGCTGGAAGGAGCCTGATCGACCTGCTCCCTCAAGCTCAG 1250
 Qy 1500 ACTGTTTCCAGCTTTGTTTTCCTTCCATGCTCTCATGCTCTTCATGTTGCTCCCTCTA 1559
 Db 1251 CGTTTCTCCGATGTTCTCTGCTCTGACCATCTTCTGTTGCTGCTGTTCCACATTTA 1310
 Qy 1560 TTCGACCCATTTAGTACAGGATTTGGCTTCTCATCACTCTGACTGAGTCCCTGGGTA 1619
 Db 1311 CAGTCACTACTCAACTCCCTCATCGGCATTGGCCATTGSCCTCTCAGGCTGCTTTTA 1370
 Qy 1620 TTATCTCTTTATTATA 1635
 Db 1371 CTCTCTCATCATCAGA 1386

RESULT 9

AAAL1955

ID AAAL1955 standard; cDNA; 1656 BP.

AC AAAL1955;

DT 01-AUG-2000 (first entry)

DE Human monp-2 cDNA #2.

KW Human; monp-2; cell proliferation; drug development; immune disease;
 KW drug development; cancer; ss.

OS Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 46..1581
 FT /*tag= a
 FT /product= "monp-2"

XX WO200017238-A1.

XX 30-MAR-2000.

XX 22-SEP-1999; 99WO-JP05205.

XX 24-SEP-1998; 98JP-0270356.

XX (TAIS) TAISHO PHARM CO LTD.

XX Yoshimoto M, Takayama K;

XX WPI; 2000-283547/24.

DR P-PSDB; AAY87630.
 XX Novel gene monp-2 encoding monocyte-specific protein with cell
 PT proliferation promoting effect, useful in studying its expression and
 PT immune function, and in developing drugs for e.g. immune diseases and
 PT cancers
 XX Example 1; Page 28-31; 34pp; Japanese.
 PS This invention describes a novel protein MONP-2 (I) and its encoding
 XX gene monp-2 (II) with a cell proliferation promoting effect isolated
 CC from a human peripheral blood. The new protein (I) and gene are useful
 CC in studying the expression abnormality and dysfunction of (I) as well
 CC as maintenance of immune function and development of drugs for e.g.
 CC immune diseases and cancers. This sequence encodes the MONP-2 protein
 CC described in the method of the invention.
 XX Sequence 1656 BP; 335 A; 445 C; 400 G; 475 T; 1 other;

Query Match 16.4%; Score 366.4; DB 21; Length 1656;
 Best Local Similarity 54.6%; Pred. No. 3.9e-86;
 Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;
 Qy 300 CGCCTTTTTCAGGAAGAGAGCGCCTTTTTCAGGAAGAGAAAGTGCAGCTGAAGAGGAAAGT 359
 Db 96 CTCGCCCTTTGGGTGATGGGGCCAGCCAGCGGGCGGAGCAGGTGAAGCTGAAGAAGAGAT 155
 Qy 360 CACTTTTACTCAGGGGAGTCTCCATTATCATTTGCCACCATCATTTGGAGCAGGAATCTTCAT 419
 Db 156 CTCACGTCTTAACGGGTGTGCTGATTTGGGGAAACATGATGGGCTCAGGCATCTTTGT 215
 Qy 420 CTCTCTTAAGGGCGTGTCTCCAGAACACGGGCGAGCTGGGCATGTCTCTCACCATCTGGAC 479
 Db 216 TTCCCCCAAGGGTGTGCTCATATACAGTGCCTTTTGGTCTCTCTGTCATCTGGGC 275
 Qy 480 GTGTGTGGGGTCTCTCCTACTATTGGAGCTTTGTCTTATGTCTGAATTTGGGAACAATAT 539
 Db 276 TGTGCGGGGCTCTCTCCGTCTTTGGGGCCCTTTGTTATGCGAACTGGCACCCACCAT 335
 Qy 540 AAGAAATCTGGAGGTCAATACACATATATTTGGAAAGTCTTTGGTCCATTACCAGCTTT 599
 Db 336 TAAGAAATCTGGGGCCAGCTATGCTTATCTGAGGGCTTTGGAGGATTTCTTCTTTT 395
 Qy 600 TGTACGAGTCTGGTGGAACTCTCTCATATACGCCCTCGAGCTACTGCTGTGATATCCCT 659
 Db 396 CATCAGACTTGGACCTCCCTGCTCATATTGAGCCACACAGCCAGGCCATCATTTGCCAT 455
 Qy 660 GGCATTGGACGCTACATTTCTGGAACCAATTTTATTCATATGTGAATCCCTGAATTCG 719
 Db 456 CACCTTTGGCAACTACATGTCACAGCTCTCTTCCCGAGCTGCTTCCGCCCTTATGCTGC 515
 Qy 720 GATCAAGCTCATACAGCTGTGGGCATACTGTAGTGTAGTGGCTTAAATAGCATGATGT 779
 Db 516 CAGCCCGCTGCTGGCTGCTGCTGCTCTCTTAACTTCAATTAACCTGTCGCTATGT 575
 Qy 780 CAGCTGAGCGCCCGCATCCAGATTTTCTTAACCTTTTCAAGCTCACAGCAATCTGAT 839
 Db 576 CAATGGGAACCTCTGTTACAGATATTTTCACTATGCTAAAGTATTTGGACTGATCGC 635
 Qy 840 AATTATAGTCCCTGGAGTTATGTCAGCTAATTAAGGTCAAAGCAGAACTTTAAAGACGC 899
 Db 636 GGTCATCGTTGCAGGCATTTTAGACTTGGCCAGGGAGCCTCTACTCAITTTTGAAGATTC 695
 Qy 900 CTTTTCAGGAAGAGATTCAGATATTACGGGGTTGCCACTGGCTTTTATATATGGAATGTA 959
 Db 696 CTTTGGGGTTTATCATTTTGCAGTGGGTGACATTTGCCCTGGCACTGTACTCAGCTGT 755
 Qy 960 TGCATATGCTGGCTGCTTTTACTCAACTTTTGTACTGAAGAGTAGAAACCCCTGAATA 1019
 Db 756 CTCCTACTCAGGCTGGACACCTCAACTATGTCACTGAAGAGATCAAGAATCCTCGAGAG 815
 Qy 1020 AACCATTCGCCCTTGCATATGATATATCCATTTGTCACCATTTGCGTATGTGCTGAC 1079

PT proliferation promoting effect, useful in studying its expression and
PT immune function, and in developing drugs for e.g. immune diseases and
PT cancers
XX
PS Claim 3; Page 26-27; 34pp; Japanese.
XX This invention describes a novel protein MONP-2 (I) and its encoding
CC gene monp-2 (II) with a cell proliferation promoting effect isolated
CC from a human peripheral blood. The new protein (I) and gene are useful
CC in studying the expression abnormality and dysfunction of (I) as well
CC as maintenance of immune function and development of drugs for e.g.
CC immune diseases and cancers. This sequence encodes the MONP-2 protein
CC described in the method of the invention.
XX
SQ Sequence 1533 BP; 301 A; 420 C; 363 G; 448 T; 1 other;
Query Match 16.4%; Score 366.4; DB 21; Length 1533;
Best Local Similarity 54.6%; Pred. No. 3.8e-86;
Matches 730; Conservative 0; Mismatches 606; Indels 0; Gaps 0;
Qy 300 CGCCCTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAGTCAGCTGAAGAGGAAAGT 359
Db 51 CTCCTCTTTGGGTGATGGGCCACGCCAGCGCGGAGCAGTGAAGCTGAAGAGGAGAT 110
Qy 360 CACTTTTACTAGGGAGTCTCCATTATCATTTGGCACCACATTTGGAGCAGGAATCTTCAT 419
Db 111 CTCACCTGCTTAACGGCGTGTCCCTGATTTGGGGAACATGATCGCTCAGGCATCTTTGT 170
Qy 420 CTCCTCTTAAGGCGTGTCTCCAGAACACGGCAGCGTGGCATGCTCTGACCATCTGCAC 479
Db 171 TTCCCCCAAGGGTGTGCTCATATACAGTGCCCTCTTTGGTCTCTCTGTCATCTGGC 230
Qy 480 GGTGTGTGGGCTCTGTCTACATTTTGGAGCTTTTCTTCTTCTGCTGAATTTGGGAACAATAT 539
Db 231 TGTGGGGGCTCTTCCGCTCTTTGGGCGCTTTGTTATGCGCAACTGGGACACCAT 290
Qy 540 AAAGAAATCTGGAGTCAATACATATATTTTGGAGTCTTTTGGTCCATTTACCAAGCTTT 599
Db 291 TAAGAAATCTGGGCGCAGCTATGCTATATCTGGAGGCTTTGGAGGATTTCTTGTGTTT 350
Qy 600 TGTAGAGTCTGGTGGAACTCTCATAAATACGCCCTGACGCTACTGCTGTGATATCCCT 659
Db 351 CATCAGACTTGGACCTCCCTGCTCATCTATTGAGCCACAGCCAGCCACATTTGCCAT 410
Qy 660 GGCAATTTGGAGCTACATTTCTGGAACCAATTTTATTAATCAATGTGAATTCCTGAACTTGC 719
Db 411 CACCTTTGCCAACTACATGTTACAGCTCTCTTCCGAGCTGCTTGGCCCTTATGCTGTC 470
Qy 720 GATCAAGCTCATACAGCTGTGGCATAACTGTAGTGTGCTCTTAATAGCATGAGTGT 779
Db 471 CAGCCGCTGTGGCTGTGCTGCTGCTCTCTTAACCTTTTTCAGCTCAGCAATTTCTGAT 530
Qy 780 CAGCTGGAGCGCCCGATCCAGATTTTCTTAACCTTTTTCAGCTCAGCAATTTCTGAT 839
Db 531 CAAATGGGAACCTTGTGTACAGATATTTTCACTATCTCTAAAGTATTGGCAGCTGATCGC 590
Qy 840 AATTATAGTCTCGAGTATGACAGCTAATTAAGGTCAAACGCAAGACTTTAAAGACGC 899
Db 591 GGTATCTGTGCGCATTTGTAGACTTGGCCAGGAGGCTCTACTCATTTTGGAGATTC 650
Qy 900 CTTTTCAGGAAGAGATTTCAAGTATTACGGGTGGCTGCTCTTTTATTTATGGAATGTA 959
Db 651 CTTTGGAGTTTCATTTTGGAGTGGGTGACATTTGCCCTGGCAGCTGACTCAGCTCTGTT 710
Qy 960 TGCATATGCTGGCTGGTGTTCCTCAACTTTTGTACTGAAGAGTAGAAGACCTTGAAAA 1019
Db 711 CTCCTACTCAGGCTGGGACACCTTCAACTATGTCTACTGAAGAGATCAAGAAATCTCTGAG 770
Qy 1020 AACCATTCCTTGCATATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 1079
Db 771 GAACCTGCCCCCTCTCCATTTGGCATCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 830
Qy 1080 AAATGTGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCTTCAATGAGCTGGC 1139

Qy 1106 ATGCTGAGAGAGTGTCTGCTTTCAATGAGTGGCAGTGAACCTTTTCTGAGCGGCTACTGG 1165
Db 967 CCACCGAGCAGATGCTGCTCGTCCGAGCGGTGGCGTGGACTTGGGAACTATCACCTGG 1026
Qy 1166 GAAATTTCTCATTAGCAGTTCGGATCTTTGTGGCTCTCTGCTGCTTGGCTCCATGAACG 1225
Db 1027 GGGTATGTCCTGGATCATCCCGTCTTCGTGGGCTGCTGCTTGGCTCCGCTCAATG 1086
Qy 1226 GTGGTGTGTTGCTGCTCAGGTTTATTTCTATTTGCTGCTGAGAGGCTCACCCTTCAG 1285
Db 1087 GGTCCCTGTTCACATCTCCAGGCTCTTCTGTTGGGTCCCGGAGGCCACCTGCGCT 1146
Qy 1286 AAATCTCTCCATGATCATGTCGCAAGACACATCTCTTACCAGCTGTTATTTGCTG 1345
Db 1147 CCATCTCTTCCATGATCCACCCACACAGCTCTCACCCCGTGGCGTCTGTTTCAGCT 1206
Qy 1346 ACCCTTTGACAATGATGCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCCCTCA 1405
Db 1207 GTGTGATGAGCTGTCTAGGCTTCTTCAAGGACATCTTCTCCGTCATCACTTCTTCA 1266
Qy 1406 GTTTTCCAGGTTGCTTTTATTTGGGCTGGCAGTTTGTGGGCTGATTTATCTTCGATACA 1465
Db 1267 GCTTCTTCAACTGGCTCTGCTGGGCTTGGCCATCATGGCATGATCTGGCTGCGCCACA 1326
Qy 1466 AATGCCACAGATGATGCTGCTTCAAGGTGCCACTGTTTATCCAGCTTTGTTTCTCT 1525
Db 1327 GAAAGCTGAGCTGTAGCGGCCCATCAAGGTGAACCTTGCCCTGCTGTTGTTCTTCACTC 1386
Qy 1526 TCACATGCTCTTCATGTTGGCTTTCCTCTATTTCCGACCCATTTAGTACAGGAGTTG 1585
Db 1387 TGGCTTGCCTTCTCTGATCGCGTCTCTTCTGGAAGACACCCGCTGGAGTGTGGCATCG 1446
Qy 1586 GCTTCTGATCACTGCTGAGTGGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1645
Db 1447 GCTTCAACATCATCTCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
Qy 1646 AACCCAGGTGGTT 1658
Db 1507 AGCCCAAGTGGCT 1519
RESULT 8
AA11954
ID AA11954 standard; cDNA; 1533 BP.
XX
AC AA11954;
XX
XX 01-AUG-2000 (first entry)
XX Human monp-2 cDNA.
XX
XX Human; monp-2; cell proliferation; drug development; immune disease;
KW drug development; cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200017238-A1.
XX
XX 30-MAR-2000.
XX
XX 22-SEP-1999; 99WO-JP05205.
XX
XX 24-SEP-1998; 98JP-0270356.
XX
XX (TAIS) TAISHO PHARM CO LTD.
XX
XX Yoshimoto M, Takayama K;
XX
XX WPI: 2000-283547/24.
XX
XX P-PSDB; AAY87630.
XX
XX Novel gene monp-2 encoding monocyte-specific protein with cell

PF	28-JAN-1999;	99WO-US01619.
XX	03-APR-1998;	98US-0080666.
PR	28-JAN-1998;	98US-0072910.
PR	24-FEB-1998;	98US-0075954.
PR	31-MAR-1998;	98US-0080114.
XX	03-APR-1998;	98US-0080515.
PA	(CHIR) CHIRON CORP.	
PA	(HYSE-) HYSEQ INC.	
XX	Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;	
PI	Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;	
PI	Janson WL, Kassam A, Kennedy GC, Kita D, Labat I;	
PI	Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;	
PI	Stache-Crain B, Sudduth-Klinger J, Williams LT;	
XX	WPI; 1999-494092/41.	
DR	Novel human genes and their expression products which are	
XX	differentially expressed in different cell types	
PT	Claim 1; Page 2164-2165; 2479pp; English.	
XX	The present invention describes a library of human polynucleotides	
CC	comprising the sequences given in AAZ12532 to AAZ17779. Also described is	
CC	a method of detecting differentially expressed genes correlated with the	
CC	cancerous state of a mammalian cell, comprising detecting at least one	
CC	differentially expressed gene product in a test sample from a cell	
CC	suspected of being cancerous, where the gene product is encoded by one	
CC	of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The	
CC	polynucleotides can be used as a source of primers and probes, which can	
CC	be used for a variety of purpose, e.g. detection of expression analysis and	
CC	mapping, tissue typing or profiling, forensic, genetic analysis and	
CC	detection of polymorphisms. Polypeptides encoded by the polynucleotides	
CC	can be used for raising antibodies for experimental, diagnostic and	
CC	therapeutic purposes. The polynucleotides may also be used to construct	
CC	arrays for diagnostics (which may be used to determine function of an	
CC	encoded protein); and to detect differences in expression levels between	
CC	two cells (e.g. to identify abnormal or diseased tissue in a human, to	
CC	identify a genetic predisposition or susceptibility to a disease such as	
CC	cancer). The polynucleotides of the invention are especially used in the	
CC	diagnosis, prognosis and management of colorectal cancer, breast cancer,	
CC	and lung cancer. The polynucleotides can also be used to screen for	
XX	peptide analogues and antagonists.	
SQ	Sequence 761 BP; 160 A; 162 C; 150 G; 250 T; 39 other;	
Query Match 23.3%; Score 520.6; DB 20; Length 761;		
Best Local Similarity 88.2%; Pred. No. 5.6e-127;		
Matches 641; Conservative 0; Mismatches 72; Indels 14; Gaps 8;		
QY	805	TTCTTAACCTTTTGAAGCTACAGCAATTTCTGATAATTATAGTCCCTGGAGTTATGCG 864
Db	18	TTTGGATCTNTGCGAGGATCCCATCGATTC--GTAATTATANNCCCTGGAGTTATGCG 75
QY	865	CTAATTAAAGGTCAACGAGCACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATT 924
Db	76	CTAATTAAAGGTCAACGCACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATT 135
QY	925	ACGCGGTTGCCACTGGCTTTTATTATGGAATGATGATGCTGGCTGGTTTACCTC 984
Db	136	ACGCGGTTGCCACTGGCTTTTATTATGGAATGATGATGCTGGCTGGTTTACCTC 195
QY	985	AACCTTTGTTACTGAAGAAGTAGAAAACCTTCCCTTGGCAATATGTATA 1044
Db	196	AACCTTTGTTACTGAAGAAGTAGAAAACCTTCCCTTGGCAATATGTATA 255
QY	1045	TCCATGCCATTGTCACCATTTGGCTATGCTGACAAATGTCGGCTACTTTACGACCAATT 1104
Db	256	TCCATGCCATTGTCACCATTTGGCTATGCTGACAAATGTCGGCTACTTTACGACCAATT 315
QY	1105	AATGCTGAGGAGTGTGCTTTTCAATGCA-GTGGCAGTGACCTTTCTGAGCGGCTACT 1163

QY 1072 GTGCTGACAAATGTCGCTTACTTTACGACCATTAATGCTGAGAGCTGCTGCTTTCAAAAT 1131
 Db 906 GTGCTGACAAATGTCGCTTACTTTACGACCATTAATGCTGAGAGCTGCTGCTTTCAAAAT 965
 QY 1132 GCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCCGATC 1191
 Db 966 GCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTTCCGATC 1025
 QY 1192 TTTGTTGGCCTCTCTGCTGCTTTGGCTCCATGAACGGTGGTGTGTGCTGCTCCAGGTTA 1251
 Db 1026 TTTGTTGGCCTCTCTGCTGCTTTGGCTCCATGAACGGTGGTGTGTGCTGCTCCAGGTTA 1085
 QY 1252 TTTCTATGTTGGCTCTCGAGAGGTCACCTTCCGAAATCTCTCCATGATTCATGTCGCGC 1311
 Db 1086 TTTCTATGTTGGCTCTCGAGAGGTCACCTTCCGAAATCTCTCCATGATTCATGTCGCGC 1145
 QY 1312 AAGCACACCTCTTACCAGCTGTTTATTTGTTGACCCCTTTGACAAATGATGATGCTTTC 1371
 Db 1146 AAGCACACCTCTTACCAGCTGTTTATTTGTTGACCCCTTTGACAAATGATGATGCTTTC 1205
 QY 1372 TCTGAGACCTCGACAGCTTTTGAATTTCTCAGTTTCCAGGTTGCGCTTTTATTTGGG 1431
 Db 1206 TCTGAGACCTCGACAGCTTTTGAATTTCTCAGTTTCCAGGTTGCGCTTTTATTTGGG 1265
 QY 1432 CTGSCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATGCTGCTTTC 1491
 Db 1266 CTGSCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATGCTGCTTTC 1325
 QY 1492 AAGTGGCAGCTGTTTATCCAGCTTTGTTTCTTCCATGCTCTTATGTTGCCCTT 1551
 Db 1326 AAGTGGCAGCTGTTTATCCAGCTTTGTTTCTTCCATGCTCTTATGTTGCCCTT 1385
 QY 1552 TCCCTCTATTGCGACCCATTTAGTACAGGATTTGGCTTCTCATCTACTGCTGCTGAGTC 1611
 Db 1386 TCCCTCTATTGCGACCCATTTAGTACAGGATTTGGCTTCTCATCTACTGCTGCTGAGTC 1445
 QY 1612 CTTGCTGATTTATCTTTATATATGCGGACAGAAACCCAGTGTTTGAATAATGTCFA 1671
 Db 1446 CTTGCTGATTTATCTTTATATATGCGGACAGAAACCCAGTGTTTGAATAATGTCFA 1505
 QY 1672 GAGAAATAACCAAGACATTTACAAATTAATCTGGAAGTTGTACCAAGAGAGATAGTTA 1731
 Db 1506 GAGAAATAACCAAGACATTTACAAATTAATCTGGAAGTTGTACCAAGAGAGATAGTTA 1565
 QY 1732 TGAATTAATGGACTTGCAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTT 1791
 Db 1566 TGAATTAATGGACTTGCAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGGATTT 1625
 QY 1792 TTACTTCAATTTCTGAAAGCTTAGAGATTTACAACTTTTGGTGATAACAAAGAGGTCAG 1851
 Db 1626 TTACTTCAATTTCTGAAAGCTTAGAGATTTACAACTTTTGGTGATAACAAAGAGGTCAG 1685
 QY 1852 TTATTTTATTTATTTAGCATATTTAGCATATTTAGCATATTTCTAAGAAATTTAGTTATAAC 1911
 Db 1686 TTATTTTATTTATTTAGCATATTTAGCATATTTAGCATATTTCTAAGAAATTTAGTTATAAC 1745
 QY 1912 TCTATGATTTATGAAAGTGAATATGAGTATTTCTATGATGTCGCAATTTCTTGAGTC 1971
 Db 1746 TCTATGATTTATGAAAGTGAATATGAGTATTTCTATGATGTCGCAATTTCTTGAGTC 1805
 QY 1972 TCTGATACCTACCTATTGGGTTAGGAGAAAGACTAGCAATTTACTATGTTGTCATTCT 2031
 Db 1806 TCTGATACCTACCTATTGGGTTAGGAGAAAGACTAGCAATTTACTATGTTGTCATTCT 1865
 QY 2032 CTACAACATATGTTAGCAGCGCAAGACCTTCAAAATTTGAGACTGAGATTTTCTGTAT 2091
 Db 1866 CTACAACATATGTTAGCAGCGCAAGACCTTCAAAATTTGAGACTGAGATTTTCTGTAT 1925
 QY 2092 ATATGGGTTTTGTAAGATGTTTTTACACACTA 2124
 Db 1926 ATATGGGTTTTGGAAGATGTTTTTACACACTA 1958

RESULT 5
 AAZ16528
 ID AAZ16528 standard; cDNA; 772 BP.
 XX
 AC AAZ16528;
 XX
 DT 12-OCT-1999 (first entry)
 XX
 XX Human gene expression product cDNA sequence SEQ ID NO:3998.
 XX
 KW Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9938972-A2.
 XX
 PD 05-AUG-1999.
 XX
 XX 28-JAN-1999; 99WO-US01619.
 XX
 PF 03-APR-1998; 98US-0080666.
 PR 28-JAN-1998; 98US-0072910.
 PR 24-FEB-1998; 98US-0075954.
 PR 31-MAR-1998; 98US-0080114.
 PR 03-APR-1998; 98US-0080515.
 XX
 XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 XX WPI; 1999-494092/41.
 XX
 PT Novel human genes and their expression products which are
 PT differentially expressed in different cell types
 XX
 PS Claim 1; Page 1897; 2479pp; English.
 XX
 CC The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
 CC polynucleotides can be used as a source of primers and probes, which can
 CC be used for a variety of purpose, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensics, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and
 CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels of an
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
 CC identify a genetic predisposition or susceptibility to a disease such as
 CC cancer). The polynucleotides of the invention are especially used in the
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
 CC and lung cancer. The polynucleotides can also be used to screen for
 CC peptide analogues and antagonists.
 XX
 SQ Sequence 772 BP; 175 A; 183 C; 158 G; 248 T; 8 other;

Query Match 28.8%; Score 643.8; DB 20; Length 772;
 Best Local Similarity 98.5%; Pred. No. 1.8e-159;
 Matches 711; Conservative 0; Mismatches 5; Indels 6; Gaps 6;

CC	forensics, gene mapping, identification of mutations, to assess
CC	biodiversity and for nutritional purposes. The present sequence is a cDNA
CC	of the invention.
XX	
SQ	Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;
Query Match	81.4%; Score 1822.2; DB 22; Length 1958;
Best Local Similarity	95.4%; Pred. No. 0;
Matches 1920; Conservative	0; Mismatches 38; Indels 55; Gaps 2;
QY	112 CAATTGTCGATAGATTTTATCATATCTGGATTTTGGATTTTCTTCATCA 171
DB	1 CAATTCGGGTCGAGATTCGTTTTCCTCTGTGTTTATTTTCCCGCTGTCCCTA 60
QY	172 CTGATTCAGGAAGCCTGTTGTCCACATCTCCAAAGAGGTTTACCTGCAGGAAAT 231
DB	61 CTATGGTCA-GAAGAGCCTGTTGTCCACATCTCCAAAGAGGTTTACCTGCAGGAAAT 119
QY	232 GTTAACGGGAGGCTGCCCTCCCTGGGCAACAAGAGCCACCTGGGCGAGCCCTTTTCA 291
DB	120 GTTAACGGGAGGCTGCCCTCCCTGGGCAACAAGAGCCACCTGGGCGAG----- 167
QY	292 GGAAGACAGCCCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAG 351
DB	168 -----GAGAAAGTGCAGCTGAAG 185
QY	352 AGGAAAGTCACTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCATCATTTGGAGCAGGA 411
DB	186 AGGAAAGTCACTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCATCATTTGGAGCAGGA 245
QY	412 ATCTTCATCTCTCTAAGGGGCTCTCCAGAACACGGGCGAGCTGGGCGATGCTCTGACC 471
DB	246 ATCTTCATCTCTCTAAGGGGCTCTCCAGAACACGGGCGAGCTGGGCGATGCTCTGACC 305
QY	472 ATCTGGACGGTGTGTGGGGTCTCTCTCACTATTTGGAGCTTTGTCTTATGCTGAATGGGA 531
DB	306 ATCTGGACGGTGTGTGGGGTCTCTCTCACTATTTGGAGCTTTGTCTTATGCTGAATGGGA 365
QY	532 ACACTATAAAGAAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTA 591
DB	366 ACACTATAAAGAAATCTGGAGGTCATTACACATATATTTTGGAGTCTTTGGTCCATTA 425
QY	592 CCAGCTTTGTACGAGTCTGGGTGGAATCTCTCATATATACGCCCTGCAGCTACTGCTGTG 651
DB	426 CCAGCTTTGTACGAGTCTGGGTGGAATCTCTCATATATACGCCCTGCAGCTACTGCTGTG 485
QY	652 ATATCCCTGGCATTTTGGACGCTACATTTCTGGAACCATTTTTTATCAATGTGAATCCCT 711
DB	486 ATATCCCTGGCATTTTGGACGCTACATTTCTGGAACCATTTTTTATCAATGTGAATCCCT 545
QY	712 GAACCTGGCATCAAGCTCATTTACAGCTGTGGGCATACCTGTAGTGTCTTAAATAGC 771
DB	546 GAACCTGGCATCAAGCTCATTTACAGCTGTGGGCATACCTGTAGTGTCTTAAATAGC 605
QY	772 ATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACTTTTGAAGCTCACAGCA 831
DB	606 ATGAGTGTGAGCTGGAGCGCCGGATCCAGATTTTCTTAACTTTTGAAGCTCACAGCA 665
QY	832 ATCTCATATATATAGTCTCCCTGGAGTATGACGCTAATTAAGGTCACCAACGAGACTTT 891
DB	666 ATCTCATATATATAGTCTCCCTGGAGTATGACGCTAATTAAGGTCACCAACGAGACTTT 725
QY	892 AAAGACGCTTTTTCAGGAAGAGATTCAGATTTTACGCGGTTGCCACTGGCTTTTATTTAT 951
DB	726 AAAGACGCTTTTTCAGGAAGAGATTCAGATTTTACGCGGTTGCCACTGGCTTTTATTTAT 785
QY	952 GGAATGTATGATATGCTGGCTTTTACCTCAACTTTGTTTACTGAGAGAGTGAAGAAC 1011
DB	786 GGAATGTATGATATGCTGGCTTTTACCTCAACTTTGTTTACTGAGAGAGTGAAGAAC 845
QY	1012 CCTGAAAAAACCAATTCCTCCCTTGCATATGATATATCCATGCGCATTTGTCACCATTTGGCTAT 1071
DB	846 CCTGAAAAAACCAATTCCTCCCTTGCATATGATATATCCATGCGCATTTGTCACCATTTGGCTAT 905

DB	1566 TGAACCTAATGCACTTGAGATCTTGGCAATCTGCCAAGGGAGAGACACAAATAGGGATT 1625
QY	1792 TTACTTTCATTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGTATATAAACAAGAGGAGTCAG 1851
DB	1626 TTACTTTCATTTCTGAAAGTCTAGAGAAATTAACAATTTGGTGTATATAAACAAGAGGAGTCAG 1685
QY	1852 TTATTTTATTCATATATTTTAGCATATTCGAACCTAATTTCTAAGAAATTTAGTTATTAAC 1911
DB	1686 TTATTTTATTCATATATTTTAGCATATTCGAACCTAATTTCTAAGAAATTTAGTTATTAAC 1745
QY	1912 TCTATGATAGTATAGAAAGTGAATATGAGTATTTCTATGAGTCCGACAAATTTCTTGAGTC 1971
DB	1746 TCTATGATAGTATAGAAAGTGAATATGAGTATTTCTATGAGTCCGACAAATTTCTTGAGTC 1805
QY	1972 TCTATGATAGTATAGAAAGTGAATATGAGTATTTCTATGAGTCCGACAAATTTCTTGAGTC 2031
DB	1806 TCTATGATAGTATAGAAAGTGAATATGAGTATTTCTATGAGTCCGACAAATTTCTTGAGTC 1865
QY	2032 CTACAACATATGTTAGCAGCGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTAT 2091
DB	1866 CTACAACATATGTTAGCAGCGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTAT 1925
QY	2092 ATATGGGTTTGTAAAGATGTTTTCACACACTA 2124
DB	1926 ATATGGGTTTGTAAAGATGTTTTCACACACTA 1958
RESULT 4	
AAH98573	
ID	AAH98573 standard; cDNA; 1958 BP.
AC	AAH98573;
XX	
DT	12-OCT-2001 (first entry)
DE	Human EST-derived coding sequence SEQ ID NO: 430.
XX	
KW	Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW	tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW	diagnostics; forensic test; gene mapping; genetic disorder;
KW	biodiversity; gene therapy; nutrition; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200154477-A2.
XX	
PD	02-AUG-2001.
XX	
PF	25-JAN-2001; 2001WO-US02687.
XX	
PR	25-JAN-2000; 2000US-0491404.
PR	17-JUL-2000; 2000US-0611746.
PR	03-AUG-2000; 2000US-0631451.
PR	15-SEP-2000; 2000US-0663870.
XX	
PA	(HYSB-) HYSEO INC.
XX	
PI	Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI	Cao Y, Drmanac RA, Zhang J, Wehrman T;
XX	
DR	WPI; 2001-476164/51.
DR	P-PSDB; AAM23914.
XX	
PT	Isolated polypeptide for treatment of diseases, diagnostics, raising
XX	antibodies and research use
PS	Claim 1; Page 484; 1275pp; English.
CC	The present invention provides the protein and coding sequences of novel
CC	proteins from a variety of organisms, including human, dog, cat, horse,
CC	cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC	urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC	from the organism of interest. They can be used in diagnostics,

giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention.

xx Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;

Query Match	81.4%	Score 1822.2;	DB 22;	Length 1958;
Best Local Similarity	95.4%	Pred. No. 0;		
Matches 1920;	Conservative	0;	Mismatches 38;	Indels 55; Gaps 2;
QY 112	CAATTGTCGATAGATTTATCATATCTCGGATTTTGGATTCCTTTGTTTCTCATCA	171		
DB 1	CAATTCGGGTCGACGATTCGTTTCCCTGTTTATTTTCCCGGCTGCCCTA	60		
QY 172	CTGGATTTCAGGAAGCCTGTGTGTCACCAATCTCCAAAGAGAGTTACCTCAGGGAAT	231		
DB 61	CTATGTCTCA-GAAAGCCTGTGTGTCACCAATCTCCAAAGAGAGTTACCTCAGGGAAT	119		
QY 232	GTAAACGGGAGCTGCTTCCCTGGCAACAGAGGACCACTGGGAGGAGCGCTTTTCA	291		
DB 120	GTAAACGGGAGCTGCTTCCCTGGCAACAGAGGACCACTGGGAGGAGCGCTTTTCA	167		
QY 292	GGAAAGAGCGCTTTTTCAGGAAGAGACGCGCTTTTCAGGAAGAGAGAAAGTCAAG	351		
DB 168	-----GAGAAAGTCAAGTCAAG	185		
QY 352	AGGAAGTCACTTTACTGAGGGAGTCTCCATTAATCAATGTCACCATCATTTGGACAGGA	411		
DB 186	AGGAAGTCACTTTACTGAGGGAGTCTCCATTAATCAATGTCACCATCATTTGGACAGGA	245		
QY 412	ATCTTCATCTCTCTTAAGGGCGTGTCCAGAACACGGGAGCGTGGGCATGCTCTGACC	471		
DB 246	ATCTTCATCTCTCTTAAGGGCGTGTCCAGAACACGGGAGCGTGGGCATGCTCTGACC	305		
QY 472	ATCTGGAGCGGTGTGGGGTCTGTCTACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGA	531		
DB 306	ATCTGGAGCGGTGTGGGGTCTGTCTACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGA	365		
QY 532	ACAACATATAAAGAAATCTGGAGGTCATACACATATATTTTGAAGTCTTTGGTCCATTA	591		
DB 366	ACAACATATAAAGAAATCTGGAGGTCATACACATATATTTTGAAGTCTTTGGTCCATTA	425		
QY 592	CCAGCTTTGTACGAGTCTGGGTGGAACCTCTCATATACGCCCTCGAGCTACTGCTGNG	651		
DB 426	CCAGCTTTGTACGAGTCTGGGTGGAACCTCTCATATACGCCCTCGAGCTACTGCTGNG	485		
QY 652	ATATCCCTGGCATTTGGACGCTACATCTCGGAACCAATTTTATTCATGTGAAATCCCT	711		

DB 486	ATATCCCTGGCATTTGGACGCTACATCTCTGAACCAATTTTATTCATCAATGTGAATCCCT	545		
QY 712	GAACCTTGGCATCAAGCTCATTAACAGCTGTGGGATCACTAGTGTAGTGTCTTAATATAG	771		
DB 546	GAACCTTGGCATCAAGCTCATTAACAGCTGTGGGATCACTAGTGTAGTGTCTTAATATAG	605		
QY 772	ATGAGTGTAGCTGGAGCGGCCGATCCAGATTTTCTTAACCTTTTGAAGCTTCACAGCA	831		
DB 606	ATGAGTGTAGCTGGAGCGGCCGATCCAGATTTTCTTAACCTTTTGAAGCTTCACAGCA	665		
QY 832	ATTCCTGATAATTAATAGTCCCTGGAGTATGCACTTAATTAAGGTCAAAAGCGAAGCTTT	891		
DB 566	ATTCCTGATAATTAATAGTCCCTGGAGTATGCACTTAATTAAGGTCAAAAGCGAAGCTTT	725		
QY 892	AAAGACGCCCTTTTTCAGGAAGAGATTCAGATTAATTAACGCGTGTGCCACTGCTTTTAT	951		
DB 726	AAAGACGCCCTTTTTCAGGAAGAGATTCAGATTAATTAACGCGTGTGCCACTGCTTTTAT	785		
QY 952	GGATGTATGATATGCTGTGCTGGTGTTCCTCAACTTTTGTACTGAAGAGTAGAAAAC	1011		
DB 786	GGATGTATGATATGCTGTGCTGGTGTTCCTCAACTTTTGTACTGAAGAGTAGAAAAC	845		
QY 1012	CCTGAAAAAACCAATTCCTTGCATATATCATCATGGCCATTTGTCAACATTTGGCTAT	1071		
DB 846	CCTGAAAAAACCAATTCCTTGCATATATCATCATGGCCATTTGTCAACATTTGGCTAT	905		
QY 1072	GTGCTGACAAATGTGGCTACTTTAGGACATTAATGTGAGGAGCTGCTGCTTCAAT	1131		
DB 906	GTGCTGACAAATGTGGCTACTTTAGGACATTAATGTGAGGAGCTGCTGCTTCAAT	965		
QY 1132	GCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGAGTTCCGATC	1191		
DB 966	GCAGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGAGTTCCGATC	1025		
QY 1192	TTTGTGTCCTCTCTCTGCTTGGCTCCATGAACGCTGTGTGTGCTGTCTCCAGGTTA	1251		
DB 1026	TTTGTGTCCTCTCTCTGCTTGGCTCCATGAACGCTGTGTGTGCTGTCTCCAGGTTA	1085		
QY 1252	TTCTATGTTGGCTCTCGAGAGGCTCACCTTCAGAAATTCCTCTCATGATTCATGCTCCG	1311		
DB 1086	TTCTATGTTGGCTCTCGAGAGGCTCACCTTCAGAAATTCCTCTCATGATTCATGCTCCG	1145		
QY 1312	AAGCACACTCTCTTACCAGCTGTATTTGTCACCTTTTGACAAATGATGATGCTCTTC	1371		
DB 1146	AAGCACACTCTCTTACCAGCTGTATTTGTCACCTTTTGACAAATGATGATGCTCTTC	1205		
QY 1372	TCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTTCAGAGTGGCTTTTATTTGGG	1431		
DB 1206	TCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTTCAGAGTGGCTTTTATTTGGG	1265		
QY 1432	CTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCAGATATGCAATGCTCTTC	1491		
DB 1266	CTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCCAGATATGCAATGCTCTTC	1325		
QY 1492	AAGTGGCAGTCTTATCCAGCTTTTGTTCCTTTCACATGCTCTCTCATGTTGCTCTTC	1551		
DB 1326	AAGTGGCAGTCTTATCCAGCTTTTGTTCCTTTCACATGCTCTCTCATGTTGCTCTTC	1385		
QY 1552	TCCCTCTATTCGACCCATTTAGTACAGGATTTGGCTTCGTCATCATCTGACTGGAGTC	1611		
DB 1386	TCCCTCTATTCGACCCATTTAGTACAGGATTTGGCTTCGTCATCATCTGACTGGAGTC	1445		
QY 1612	CCTCGTATTTATCTCTTTATTTATATGAGCAACAAACCCAGGTTTGAATAATGTCA	1671		
DB 1446	CCTCGTATTTATCTCTTTATTTATATGAGCAACAAACCCAGGTTTGAATAATGTCA	1505		
QY 1672	GAGAAATATACAGCAACATTTACAAATTAATCTGGAAGTTGTACAGAGAAAGATTA	1731		
DB 1506	GAGAAATATACAGCAACATTTACAAATTAATCTGGAAGTTGTACAGAGAAAGATTA	1565		
QY 1732	TGAACATAGTACATGAGATCTTGGCAATCTGCCAAGGGGAGACACAAAATAGGATTT	1791		

QY	903	TTGAGGAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGAAATGATGC	962	QY	1983	CCATTTGGGTTAGGAGAAAGACTAGACAATTACTATGTGCTCATCTCTACACATAT	2042
Db	847	TTGAGGAGAGATTCAAGTATTACGGGTTGCCACTGGCTTTTATTATGAAATGATGC	906	Db	1927	CCATTTGGGTTAGGAGAAAGACTAGACAATTACTATGTGCTCATCTCTACACATAT	1986
QY	963	ATATGCTGGCTGCTTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCCCTGAAAAC	1022	QY	2043	GTTAGCACGGCAAGAACCTTCAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT	2102
Db	907	ATATGCTGGCTGCTTTTACCTCAACTTTGTTACTGAAGAAGTAGAAAACCCCTGAAAAC	966	Db	1987	GTTAGCACGGCAAGAACCTTCAAATTTGAAGACTGAGATTTTCTGTATATATGGGTTT	2046
QY	1023	CATTCCCTTGCATATGATATCCATGCCATTTGCACCATTTGCTTCAAAATGAGTGGCAGT	1082	QY	2103	GTTAAAGATGTTTACACACTACAGATGCTATCTACTGTGAAAGTGTGTTTCAATTCGAA	2162
Db	967	CATTCCCTTGCATATGATATCCATGCCATTTGCACCATTTGCTTCAAAATGAGTGGCAGT	1026	Db	2047	GTTAAAGATGTTTACACACTACAGATGCTATCTACTGTGAAAGTGTGTTTCAATTCGAA	2106
QY	1083	TGTGGCTTACTTTACGACCAATTAATGCTGAGGAGTGTGCTTTTCAAAATGAGTGGCAGT	1142	QY	2163	AAAAAGCATACATCATGATTATGGCAAGAGGAGAGAGGTA	2204
Db	1027	TGTGGCTTACTTTACGACCAATTAATGCTGAGGAGTGTGCTTTTCAAAATGAGTGGCAGT	1086	Db	2107	AAAAAGCATACATCATGATTATGGCAAGAGGAGAGAGGTA	2148
QY	1143	GACCTTTTCTGAGGGCTACTGGAAATTTCTCATTTAGCAGTTCGATCTTTGTTGCCCT	1202	RESULT 3			
Db	1087	GACCTTTTCTGAGGGCTACTGGAAATTTCTCATTTAGCAGTTCGATCTTTGTTGCCCT	1146	ABA09201 standard; cDNA; 1958 bp.			
QY	1203	CTCTGCTTTTGGCTCCATGAACGCTGGTGTGTTGCTGCTCCAGGTTATTTCTATGTTGC	1262	AC ABA09201;			
Db	1147	CTCTGCTTTTGGCTCCATGAACGCTGGTGTGTTGCTGCTCCAGGTTATTTCTATGTTGC	1206	DT			
QY	1263	GTCTCGAGAGGTCACCTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCACACTCC	1322	XX			
Db	1207	GTCTCGAGAGGTCACCTTCCAGAAATCTCTCCATGATTCATGTCGCAAGCACACTCC	1266	XX			
QY	1323	TCTACCACTGTTATTTGTTGACCCCTTTGACAATGATATCTCTCTCTGAGACCT	1382	XX			
Db	1267	TCTACCACTGTTATTTGTTGACCCCTTTGACAATGATATCTCTCTCTGAGACCT	1326	XX			
QY	1383	CGACAGCTTTTGAATTTCTCACTTTTGGAGTGCTTTTATTTGGCTGCGACTGC	1442	XX			
Db	1327	CGACAGCTTTTGAATTTCTCACTTTTGGAGTGCTTTTATTTGGCTGCGACTGC	1386	XX			
QY	1443	TGGGCTGATTATCTTCATACAAATGCCAGATATGCATGCTCTTTTCAAGTGCCACT	1502	XX			
Db	1387	TGGGCTGATTATCTTCATACAAATGCCAGATATGCATGCTCTTTTCAAGTGCCACT	1446	OS			
QY	1503	GTTCAATCCAGCTTTGTTTCTTCAATGCTCTTTCATGGTGGCTTTTCCCTCTATTC	1562	OS			
Db	1447	GTTCAATCCAGCTTTGTTTCTTCAATGCTCTTTCATGGTGGCTTTTCCCTCTATTC	1506	PN			
QY	1563	GGACCAATTTAGTACAGGATTTGCTTCTCATCTGACTGAGTCCCTGCGTATTA	1622	XX			
Db	1507	GGACCAATTTAGTACAGGATTTGCTTCTCATCTGACTGAGTCCCTGCGTATTA	1566	XX			
QY	1623	TCTCTTTTATATATGGCAAGAAACCCAGGTGTTTGAATATATGTCAGAGAAATAC	1682	PD			
Db	1567	TCTCTTTTATATATGGCAAGAAACCCAGGTGTTTGAATATATGTCAGAGAAATAC	1626	XX			
QY	1683	CAGAAATTTACAAATTAATGAAATTTTACCAGAAAGATTAAGTATGAACATATGG	1742	XX			
Db	1627	CAGAAATTTACAAATTAATGAAATTTTACCAGAAAGATTAAGTATGAACATATGG	1686	XX			
QY	1743	ACTTGAGATCTTGGCAATCTGCCAAGGGGAGACAAAATAGGATTTTACTTCAATTT	1802	XX			
Db	1687	ACTTGAGATCTTGGCAATCTGCCAAGGGGAGACAAAATAGGATTTTACTTCAATTT	1746	XX			
QY	1803	TCTGAAAGCTAGAGAAATTTACAACTTTTGTGATAACAAAGGAGTCAATTTTATTT	1862	XX			
Db	1747	TCTGAAAGCTAGAGAAATTTACAACTTTTGTGATAACAAAGGAGTCAATTTTATTT	1806	XX			
QY	1863	CATATATTTTACGATATTTGAAATTTTCTAAGAAATTTAGTTATATACTCTATGATTT	1922	XX			
Db	1807	CATATATTTTACGATATTTGAAATTTTCTAAGAAATTTAGTTATATACTCTATGATTT	1866	CC			
QY	1923	ATAGAAAGTGAATATGCAATTTTCTATGAGTGGCACAATTTCTGAGTCTCTGATACCTA	1982	CC			
Db	1867	ATAGAAAGTGAATATGCAATTTTCTATGAGTGGCACAATTTCTGAGTCTCTGATACCTA	1926	CC			

Human cystine/Glu transporter homologue cDNA, SEQ ID NO:977.

Human; cytokine; cell proliferation; cell differentiation; growth factor; hematopoiesis regulation; tissue growth; immunomodulator; activin; inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis; proliferation; metastasis; cancer; tumour; haematopoietic disorder; myeloid cell disorder; lymphoid cell disorder; asthma; arthritis; chronic inflammatory condition; proliferative retinopathy; atherosclerosis; coronary heart disease; arterial ischaemia; bone disorder; osteoporosis; vascular growth disorder; tissue regeneration; wound healing; infection; immune disorder; cell culture; drug screening; gene therapy; antiinflammatory; antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic; cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial; antifungal; vulnery; antiulcer; ss.

Homo sapiens.

WO200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US03800.

03-FEB-2000; 2000US-0496914.

27-APR-2000; 2000US-0560875.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI: 2001-457740/49.

P-PSDB; ABB11957.

Human proteins and DNA encoding sequences useful for preventing, treating or ameliorating a medical condition in a mammalian subject e.g. arthritis and cancer -

Claim 1; Page 835; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby

XX	WPI: 2001-639201/73.
XX	New human lung-specific polynucleotides and polypeptides for the
XX	diagnosis and treatment of disease e.g. lung cancer -
PT	Claim 1; Page 332; 378pp; English.
XX	The invention relates to isolated lung tumour-specific proteins and
XX	their corresponding cDNA molecules. Lung tumour-specific proteins and
CC	their antigen-presenting cells are useful for stimulating and/or
CC	expanding T cells specific for a tumour protein, and for inhibiting
CC	the development of cancer. The invention also relates to a composition
CC	useful for stimulating an immune response, and for treating cancer. The
CC	lung tumour specific oligonucleotide is useful in gene therapy and for
CC	diagnosis, detection and treatment of lung cancer. The present sequence
CC	is a cDNA encoding human lung tumour-specific protein.
XX	Sequence 2239 BP; 619 A; 444 C; 493 G; 683 T; 0 other;
XX	
QY	Query Match 100.08; Score 2239; DB 23; Length 2239;
PT	Best Local Similarity 100.08; Pred. No. 0;
PT	Matches 2239; Conservative 0; Mismatches 0; Indels 0; Gaps
XX	
QY	1 GGAGGTTGAAGTCGACGAGATCATGCCAGCGTGGTGACAGTCTGTCTCAAAC 60
DB	1 GGAGGTTGAAGTCGACGAGATCATGCCAGCGTGGTGACAGTCTGTCTCAAAC 60
QY	61 AGAATTAAAGGAAAGAAAGAAAGAGAGAGAGAGAAATTCAGAGCCAAATTGTGG 120
DB	61 AGAATTAAAGGAAAGAAAGAAAGAGAGAGAGAGAAATTCAGAGCCAAATTGTGG 120
QY	121 CATAGATTTATCATATCTCGGATTTTGGGATTCCTTTGTTTCTCATCCTGGATTCA 180
DB	121 CATAGATTTATCATATCTCGGATTTTGGGATTCCTTTGTTTCTCATCCTGGATTCA 180
QY	181 GGAAAGCCGTGTGTCTCCACATCTCCAAGGAGGTTACCTGTCAGGGAAATGTTAACGGG 240
DB	181 GGAAAGCCGTGTGTCTCCACATCTCCAAGGAGGTTACCTGTCAGGGAAATGTTAACGGG 240
QY	241 AGGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGCGAGGAGCCCTTTTCAGGAAGAGAC 300
DB	241 AGGCTGCCTTCCCTGGGCAACAAGGAGCCACCTGGGCGAGGAGCCCTTTTCAGGAAGAGAC 300
QY	301 GCCTTTTCAGGAGAGAGCGCTTTTCAGGAAGAGAGAAAGTCAGCTGAAGAGGAAAGTC 360
DB	301 GCCTTTTCAGGAGAGAGCGCTTTTCAGGAAGAGAGAAAGTCAGCTGAAGAGGAAAGTC 360
QY	361 ACTTTACTGAGGGAGTCTCCATTATCATTTGGCAACCATCATTTGGAGCAGGATCTTCATC 420
DB	361 ACTTTACTGAGGGAGTCTCCATTATCATTTGGCAACCATCATTTGGAGCAGGATCTTCATC 420
QY	421 TCTCCTTAAGGGCGTCTCCAGAACCGGACGGTGGGATGTCCTGACCAATCTGGACG 480
DB	421 TCTCCTTAAGGGCGTCTCCAGAACCGGACGGTGGGATGTCCTGACCAATCTGGACG 480
QY	481 GTGTGTGGGGTCTGTGCACTATTTGGAGCTTTGTCTATGCTGAAATGGGAACAACATA 540
DB	481 GTGTGTGGGGTCTGTGCACTATTTGGAGCTTTGTCTATGCTGAAATGGGAACAACATA 540
QY	541 AAGAAATCTGGAGGTCATTACACATATATTTGGAAGCTCTTTGGTCCATTACCACTTTT 600
DB	541 AAGAAATCTGGAGGTCATTACACATATATTTGGAAGCTCTTTGGTCCATTACCACTTTT 600
QY	601 GTACGAGTCTGGGTGGAACTCCTCAATATAGCCCTCGACGTACTGCTGTGATATCCCTG 660
DB	601 GTACGAGTCTGGGTGGAACTCCTCAATATAGCCCTCGACGTACTGCTGTGATATCCCTG 660
QY	661 GCATTTGGAGCGCTACATTTCTGGAACCATTTTTTATTCGAATGGAATCCCTGAACTTGGG 720
DB	661 GCATTTGGAGCGCTACATTTCTGGAACCATTTTTTATTCGAATGGAATCCCTGAACTTGGG 720
QY	721 ATCAAGCTCATACAGCTGGGCGATACACTGTAGTGATGGTCCCTTAATACATGAGTGTC 780

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 03:52:53 ; Search time 521 Seconds
(without alignments)
9677.974 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaaaaaa 2239

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_101002.*
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2239	100.0	2239	23	Human lung tumour-
2	2080.8	92.9	5981	23	Human lung tumour-
3	1822.2	81.4	1958	22	Human cystine/Glu
4	1822.2	81.4	1958	22	Human EST-derived
5	643.8	28.8	772	20	Human gene express
6	520.6	23.3	761	20	Human gene express
7	370.8	16.6	4539	21	Human L-type amino
8	366.4	16.4	1533	21	Human morp-2 cDNA.
9	366.4	16.4	1656	21	Human morp-2 cDNA

10	366.4	16.4	2072	20	AAZ23240	Human amino acid p
11	366.4	16.4	2072	20	AAZ86999	Human amino acid p
12	362.2	16.2	1524	21	AAA08393	Human L-type amino
13	353.6	15.8	4585	22	ABA09358	Human 4F2 light ch
14	344.8	15.4	4117	22	AAZ77276	Rat cDNA encoding
15	343.8	15.4	1632	21	AAZ55789	CDNA encoding rat
16	343.8	15.4	2155	21	AAZ00605	Human membrane tra
17	343.8	15.4	4211	22	AAZ77278	Human CDNA encodin
18	343.8	15.4	4261	21	AAZ08373	Human OREF ORF2900
19	342.2	15.3	3455	21	AAZ08373	Nucleotide sequenc
20	331.6	14.8	1716	22	AAH74767	Drosophila melanog
21	329.2	14.7	2476	23	ABL28785	Drosophila melanog
22	328.4	14.7	2090	23	ABL10465	Nucleotide sequenc
23	325	14.5	1997	23	ABL02951	Drosophila melanog
24	315.6	14.1	1918	22	AAH74769	Drosophila melanog
25	311.4	13.9	1554	23	ABL12149	Nucleotide sequenc
26	308.2	13.8	800	20	AAZ16609	Drosophila melanog
27	303	13.5	596	24	ABK27651	Human gene express
28	301	13.4	1461	22	AAF94439	Human gene express
29	301	13.4	1758	22	AAF94439	Human hydrophobic
30	300	13.4	300	20	AAZ14980	Human hydrophobic
31	300	13.4	300	20	AAZ14962	Human gene express
32	288	12.9	288	20	AAZ14304	Human gene express
33	287.8	12.9	300	21	AAA00880	Human gene express
34	286.2	12.8	3735	22	ABA09032	Human amino acid t
35	281	12.6	337	23	AAZ23462	Human lung tumour-
36	260.8	11.6	1551	23	ABL15177	Drosophila melanog
37	238.4	10.6	3984	24	ABK84083	Human cDNA differe
38	238.4	10.6	3984	24	ABL62111	Colon adenocarcino
39	238.4	10.6	3984	24	ABL62111	Breast cancer rela
40	230	10.3	575	24	ABL37718	Human colon tumour
41	229	10.2	1564	21	AAZ26346	Human secreted pro
42	226.8	10.1	2256	21	AAZ99060	Human pancreatic c
43	224	10.0	1161	23	AAZ76312	DNA encoding novel
44	215.2	9.6	1397	24	ABL90625	Human polynucleoti
45	211.4	9.4	1011	19	AAV68500	Human tumour-assoc

ALIGNMENTS

RESULT 1	AAZ23460
ID	AAZ23460 standard; cDNA; 2239 BP.
XX	AAZ23460;
AC	AAZ23460;
XX	
DT	26-FEB-2002 (first entry)
DE	Human lung tumour-specific 19A4 cDNA.
XX	
XX	Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW	antisense-therapy; vaccine; immune response; lung cancer; 19A4; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200172295-A2.
XX	
PD	04-OCT-2001.
XX	
PF	28-MAR-2001; 2001WO-US09991.
XX	
PR	29-MAR-2000; 2000US-0538037.
PR	05-JUN-2000; 2000US-0588937.
PR	18-AUG-2000; 2000US-0640878.
PR	22-SEP-2000; 2000US-234517P.
PR	01-NOV-2000; 2000US-0704512.
XX	14-DEC-2000; 2000US-0738973.
PA	(CORI-) CORIXA CORP.
XX	
PI	Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY; Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;

Matches 719; Conservative 0; Mismatches 629; Indels 10; Gaps 2;									
QY	334	GAGAAAGTGCAGCTGAGAGGAAAGTCACTTTACTGAGGGAGTCTCCATTATCATTTGGC	393	QY	1405	AGTTTTGCCAGTGGCTTTTATTGGCTGGCAGTTGCTGGCTGATTTATCTTCGATAC	1464		
Db	241	GAGCGGTGGCGCTCAAGAGGAGATCGGGCTGCTGAGCGCTGCACCATCATCATCGG	300	Db	1320	TCCTTCATCAACTACCTCTGCTACGGCGTCACAATCTGGCGCTGCTGCTGCGCTGG	1379		
QY	394	ACCATCAATGGAGAGAAATCTCATCTCTCTAAGGGCTGCTCCAGAACACGGGCAGC	453	QY	1465	AAATGCCAGATATGATCATGCTCTTTCAAGGTGCCACTGTTTCATCCAGCTTTGTTTTC	1524		
Db	301	AACATCATCGGCTCGGGCATCTTCATCTCGCCCAAGGGGTCTGGAGCACTCAGGCTCC	360	Db	1380	AGCGGCTTGCACCTCCACAGGCCCATCAAGSTGAACCTTCTCATCCCGTGGCGTACTTG	1439		
QY	454	GTGGGCATCTCTTGACCATCTGCAGCGTGTGTGGGTCTCTGCATATTTGGAGCTTTG	513	QY	1525	TTACATGCTCTTTCATGTTGGCTTTCCTCTATTTCGGACCCATTTAGTACAGGGATT	1584		
Db	361	GTGGTCTGGCCCTGTTCTGCTGGGTCTGCTGGTGGGGCTGACGGCTCTGGGCTCCCTC	420	Db	1440	GTCTTCTGGGCCTTCTCTGCTTTCAGCTTCATCTCAGAGCTATGCTGTGGGGTC	1499		
QY	514	TCATTATGCTGAATGGGAACAACATATAAGAAATCTCGAGGCTATACACATATATTTG	573	QY	1585	GGCTTCGTCATCACTCTGACTGGAGTCCCTGCTGATTTATCTTTATATATGACGACAAG	1644		
Db	421	TGCTATGACAGCTGGGAGTCGCCATCCCAAGTCTTGGCGGGACTACGGCTACGTCACA	480	Db	1500	GGGCTCATCATCATCTTACGGGGTGCCCATTTCTTCTGGGAGTGTCTTGAGAAGC	1559		
QY	574	GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGGTGGAACCTCTCTATAATACG	633	QY	1645	AAACCCAGTGGTTTAGAATAATGTCAGAGAAATAAC	1682		
Db	481	GAGATCTTCGGGGCTGGCTGGCTTCTGCTGCTCTGGAGCGCGCTCTCATCATGTAC	540	Db	1560	AAACCAAGTGTGTGCACAGACTCACAGAGTCCATGAC	1597		
QY	634	CCTGCACTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATTTCTGGAACCAATTTT	693	RESULT 11					
Db	541	CCACGAGC-CTTGCTGTCACTCCGTGACCTTCTCCAACCTACGCTGCTGAGCGCGTGT	599	BJ075046	LOCUS	403 bp	mRNA	linear	EST 11-DEC-2001
QY	694	ATTCATGTAATCCCTGAACTTTCGATCAAGCTCATTTACAGCTGTGGGCATAACTGTA	753	DEFINITION	BJ075046 NIBB Mochii normalized Xenopus tailbud library Xenopus				
Db	600	CCCAACTGCATCCCGCCACACAGCTCCCGGTGTGTCTTCCATGGCTGCGCTGATGCTC	659	ACCESSION	BJ075046				
QY	754	GTGATGGTCTTAATAGCATGAGTGCAGTCCAGTGGAGCGCCGGATCCAGATTTTCTTAAC	813	KEYWORDS	EST.				
Db	660	CTGACATGGGTGAACAGCTCCAGTGTGCGTGGCCAGCGCATCCAGGACATGTTTCA	719	SOURCE	African clawed frog.				
QY	814	TTTTTCAAGCTCACAGCAATTTCTGATTAATATAGTCCCTGGAGTTATGACGCTAAATTA	873	ORGANISM	Xenopus laevis				
Db	720	GCGGGGAAGCTGCTGGCCTTGTCCCTCATCATCGCGCTGGCGCTTCTCCAGATCTTCCA	779	REFERENCE	Xenopus laevis				
QY	874	GGTCAACACGCGAA-----CTTTAAGACGCTTTTTCAGGAAGATTTCAAGTATT	924	AUTHORS	Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara				
Db	780	GGACACTTCGAGGAGCTGAGCGCCAGCAATGCTTTGCTTCTGGATGACGCCCTCCGTG	839	TITLE	Expressed genes in X. laevis embryo				
QY	925	ACGGGTTGCCACTGGCTTTTATATGAAATGATGATATGCTGCTGCTGTTTACCTC	984	JOURNAL	Unpublished (2001)				
Db	840	GGACACCTGGCCCTGGCTTCTCCAGGCTCCTTCGCTTCACTGCTGGCTGGAATTCCTC	899	COMMENT	Contact: Tadasi Shin-i				
QY	985	AACTTTGTTACTGAAGAGTGAAGAACCTGAAAAACCATTTCCCTTGAATATGATA	1044	National Institute of Genetics					
Db	900	AACTATGTCACCGAGGAGATGTTGACGCCGGAAGAACCTACCTCGCGCATCTTCATC	959	1111 Yata, Mishima, Shizuoka 411-8540, Japan					
QY	1045	TCCATGGCCATTTGACCATTTGCTATGCTGACAAATGTGGCTACTTTACGACCAT	1104	Tel: 81-559-81-6856					
Db	960	TCCATPCCCACTGGTGACCTTGTGTACAGTTTCAACCAATTTGCTACTTTCACGGCCATG	1019	Fax: 81-559-81-6855					
QY	1105	AATGCTGAGGAGCTGCTGCTTTCAATGCAGTGGCAGTGACCTTTTCTGAGGGGCTACTG	1164	Email: tshiniegenes.nig.ac.jp.					
Db	1020	TCGCCCCAGAGCTGTCTCTCTCAATGGGTGGTGTGACCTTCGGGGAAGAGCTGCTG	1079	Location/Qualifiers					
QY	1165	GGAAATTTCTATTAGCAGTTCGATCTTTGTTGCGCTCTCTGCTTGGCTTCCATGAAC	1224	1. .403					
Db	1080	GGCTACTTTTCTGGGTATGCTGCTCCCGTGGCTGTGTCAACCTTCGGAGGATCAAT	1139	/organism="Xenopus laevis"					
QY	1225	GGTGTGTGTTGCTGTCTCCAGGTTATTTCTATGTTGGCTTCGAGAGGTCACCTTCCA	1284	/db_xref="taxon:8355"					
Db	1140	GGTTACTGTTCACCTACTCCAGGCTGTGTTCTCTGGAGCGCGGAGGGCACCTGCCC	1199	/clone="XL087e10"					
QY	1285	GAATCCTCTCCATGATTCATGTCGCAAGCACACTCTCTACACAGCTGTTTATTTTG	1344	/clone_lib="NIBB Mochii normalized Xenopus tailbud					
Db	1200	AGGCTGCTGGCCATGATCCAGCTCAGACACTGCACCCCACTCCCGCCCTCTGCTGCT	1259	library"					
QY	1345	CACCCCTTTGACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTCCTC	1404	/tissue_type="whole embryo"					
Db	1260	TGCGGGGCCACACGCGCTCATCATGCTGCTGGGCGACACGTCACAGCTCATCACTATGTG	1319	/dev_stage="stage 25"					
Query Match 11.2%; Score 250; DB 13; Length 403;									
Best Local Similarity 78.0%; Pred. No. 7.7e-38;									
Matches 312; Conservative 0; Mismatches 87; Indels 1; Gaps 1;									
QY	919	AGTATTACGCGTGTCCCACTGGCTTTTATTATGGAATGTATCATATGCTGGCTGGTTT	978						
Db	4	AGTGTATTGGGTTTACCCTTGGCATTTCTTCTGGAATGTATCATATGCTGATGTTT	63						
QY	979	TACCTCAACTTTGTTTACTGAGAAGTAGAAAACCTCGAAACCACTTCCCTTCGAATA	1038						

Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO&t2=PMO-GN0344-
271200-001-f12&t3=2000-12-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 106
High quality sequence stop: 117.
Location/Qualifiers
1. .497
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0344"
/dev_stage="Adult"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT 103 a 120 c 113 g 161 t
ORIGIN
Query Match 16.1%; Score 359.8; DB 14; Length 497;
Best Local Similarity 88.1%; Pred. No. 7.6e-59;
Matches 438; Conservative 0; Mismatches 52; Indels 7; Gaps 4;
QY 1065 TGGCTATGTCGTGCAAAATGCGCTACTTACGACCAATTAATGCTGAGAGCTGCTGCT 1124
Db 6 TTGGTATGTCGTGCAAAATGCTGCTACTTACGACCAATTAATGCTGAGAGCTGCTGCT 64
QY 1125 TTCAATGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 1184
Db 65 TCAAAAGCCGTGGCGT---AACCTTTTGTAGCCGTAATGGAAATTTCTCAATAAGCAG 121
QY 1185 TCCGATCTTTGTCGCCCTCTCCTGCTTGGCTCCATGAACGGTGGTGTGCTGCTGCTC 1244
Db 122 TCCGATCTTTG-TGGCTCTCTGCTTGGCTCCATGAACGGTGGTGTGCTGCTGCTC 180
QY 1245 CAGGTTATTTCTATGTTGGTCTGCTGAGAGGTCACCTTCCAGAAATCCTTCCATGATTC 1304
Db 181 CAGGTTATTTCTATGTTGGTCTGCTGAGAGGTCACCTTCCAGAAATCCTTCCATGATTC 240
QY 1305 TGTCCGCAAGCACACTCTCTACAGCTGTTATTTGTTTGGACCTTTGACAATGATAAT 1364
Db 241 TGTCCGCAAGCACACTCTCTACAGCTGTTATTTGTTTGGACCTTTGACAATGATAAT 300
QY 1365 GCTCTTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1424
Db 301 GCTCTTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 360
QY 1425 TATTGGG---CTGGCAGTGTGCTGCTGAGAGGTCACCTTCCAGAAATCCTTCCATGATTC 1482
Db 361 TATTGGGCGCTGCGAGATGCTGGGGTGAATTTATCTTCGATACAAATGCCAGATGATGC 420
QY 1483 CGTCCCTTCAAGTGCCACTGTTTCATCCAGCTTGTGTTTCTTCCATGCTCCTTTCATG 1542
Db 421 CGTCCCTTCAAGTGCCACTGTTTCATCCAGCTTGTGTTTCTTCCATGCTCCTTTCATG 480
QY 1543 GTTGGCCCTTTCCCTCTA 1559

FEATURES
source
TITLE
JOURNAL
MEDLINE
COMMENT
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO&t2=PMO-GN0344-
271200-001-f12&t3=2000-12-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 106
High quality sequence stop: 117.
Location/Qualifiers
1. .497
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0344"
/dev_stage="Adult"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT 131 a 50 c 69 g 115 t
ORIGIN
Query Match 16.2%; Score 363; DB 14; Length 371;
Best Local Similarity 98.6%; Pred. No. 2.1e-59;
Matches 363; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1661 GAATAATGTCAGAGAAATTAACAGACATTTACAAATAATCTGGAAGTTGTACAGAG 1720
Db 1 GAATAATGTCAGAGAAATTAACAGACATTTACAAATAATCTGGAAGTTGTACAGAG 60
QY 1721 AAGATAAGTTAAGTAATGAGCTGAGATCTTGGCAATCTGCCAAGGGGAGACACAA 1780
Db 61 AAGATAAGTTAAGTAATGAGCTGAGATCTTGGCAATCTGCCAAGGGGAGACACAA 120
QY 1781 AATAGGGAATTTTACTTCATTTCTGAAAGCTAGAGAAATTTACAACTTTGGTGATAAACA 1840
Db 121 AATAGGGAATTTTACTTCATTTCTGAAAGCTAGAGAAATTTACAACTTTGGTGATAAACA 180
QY 1841 AAGAGAGTCAGTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCTAAGAAAT 1900
Db 181 AAGAGAGTCAGTATTTTATTCATATATTTTAGCATATTCGAACTAATTTCTAAGAAAT 240
QY 1901 TTAGTTATACCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1960
Db 241 TTAGTTATACCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 1961 ATTCTTGAGCTCTCTGATACCTACCTATTTGGGGTTAGAGAAAGACTAGACAAATTA 2020
Db 301 ATTCTTGAGCTCTCTGATACCTACCTATTTGGGGTTAGAGAAAGACTAGACAAATTA 360
QY 2021 GTGGTCAT 2028
Db 361 GTGGTCAT 368
RESULT 8
LOCUS BQ370588 497 bp mRNA linear EST 21-MAY-2002
DEFINITION PMO-GN0344-271200-001-f12 GN0344 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ370588
VERSION BQ370588.1 GI:21046102
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 497)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Simpson, D.H.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

QY 378 CTCATTATCATTTGGCACCACATCTTGGACGAGAAATCTTCATCTCTCCTTAAGGGGTGCT 437
 Db |||||
 QY 309 CTCATTATCATTTGGCACCACATCTTGGACGAGAAATCTTCATCTCTCCTTAAGGGGTGCT 368
 Db |||||
 QY 438 CCAGAACACGGGAGCGTGGGCGATGCTCTGACCATCTCTGACGGGTGTGGGGTCTCTGTC 497
 Db |||||
 QY 369 CCAGAACACGGGAGCGTGGGCGATGCTCTGACCATCTCTGACGGGTGTGGGGTCTCTGTC 428
 Db |||||
 QY 498 ACTATTTGGAGCTTTGCTTTATGCTGAATTTGGGACCACTATAAAGAAATCTGGAGTCA 557
 Db |||||
 QY 429 ACTATTTGGAGCTTTGCTTTATGCTGAATTTGGGACCACTATAAAGAAATCTGGAGTCA 488
 Db |||||
 QY 558 TTACACATATATTTTGGAAAGTCTTTGGTCCATTACCACTTTTGGACAGTCTCTGGTGA 617
 Db |||||
 QY 489 TTACACATATA-TTTGGAAGTCTTTGGTCCATTACCACTTTTGGACAGTCTCTGGTGA 547
 Db |||||
 QY 618 ACTCCTCATATACGCCCTGGACGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACAT 677
 Db |||||
 QY 548 ACTCCTCATATACGCCCTGGACGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACAT 607
 Db |||||
 QY 678 TCTGGAACCATTTTATTTCAATGTAATCCCTGAACTTGGATTCGATCAAGCTCATTTACAGC 737
 Db |||||
 QY 608 TCTGGAACCATTTTATTTCAATGTAATCCCTGAACTTGGATTCGATCAAGCTCATTTACAG- 666
 Db |||||
 QY 738 TGTGGGCATACTGTAGTATGCTGCTAAATAGCATGAGTGTGAG 782
 Db |||||
 QY 667 TGTGGGGATTACTGTAGTATGCTGCTAAATAGCATGAGTGTGAG 711
 Db |||||

RESULT 6
 BG284503
 LOCUS 602408645F1 NIH_MGC_91 936 bp mRNA linear EST 21-FEB-2001
 DEFINITION mRNA sequence.
 ACCESSION BG284503
 VERSION BG284503.1 GI:13035516
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 936)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Cloning Distribution: Inocyte Genomics, Inc.
 Cloned through the I.M.A.G.E. Consortium/LLNL at:
 Plate: LLNL0463 row: 1 column: 11
 High quality sequence stop: 795.
 Location/Qualifiers
 1..936
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4537810"
 /clone_lib="NIH_MGC_91"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 248 a 202 c 251 g 235 t

FEATURES
 source
 18.4%; Score 411.2; DB 12; Length 936;
 Query Match
 BASE COUNT
 ORIGIN

Best Local Similarity 87.9%; Pred. No. 9-le-69;
 Matches 494; Conservative 0; Mismatches 13; Indels 55; Gaps 2;
 QY 182 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGTTTACTGTCAGGAAATGTTTACGGGA 241
 Db |||||
 Db 278 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGTTTACTGTCAGGAAATGTTTACGGGA 337
 Db |||||
 QY 242 GGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGCGAGGACCCCTTTTTCAGGAAGAGACG 301
 Db |||||
 Db 338 GGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGCGAG----- 375
 QY 302 CCTTTTCAGGAAGAGACGCGCTTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGAAAGTCA 361
 Db |||||
 Db 376 -----GAGAAAGTGCAGCTGAAGAGAAAGTCA 403
 QY 362 CTTTACTGAGGGAGTCTCCATTTATTTGGCAGCATTCATTGGAGGAGGATCTTCATCT 421
 Db |||||
 Db 404 CTTTACTGAGGGAGTCTCCATTTATTTGGCAGCATTCATTGGAGGAGGATCTTCATCT 463
 QY 422 CTCCTAAGGGGTGCTCCAGAACAGCGGCGAGGCGATGCTCTGACCATCTGGACGG 481
 Db |||||
 Db 464 CTCCTAAGGGGTGCTCCAGAACAGCGGCGAGGCGATGCTCTGACCATCTGGACGG 523
 QY 482 TGTGTGGGTCTCTGCTACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAAACATATA 541
 Db |||||
 Db 524 TGTGTGGGTCTCTGCTACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAAACATATA 583
 QY 542 AGAAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTGTGCTTATTCACCATCTAA 601
 Db |||||
 Db 584 AGAAATCTGGAGGTCATTACACATATATTCGGAAGTCTTTGTGCTTATTCACCATCTAA 643
 QY 602 TACGAGTCTGGTGGAACTCTCTCATATACGCGCTCGAGTCTGCTGCTGATATCCCTGG 661
 Db |||||
 Db 644 TACGAGTCTGGTGGAACTCTCTCATATACGCGCTCGAGTCTGCTGCTGATATCCCTGG 703
 QY 662 CATTGGAGCGCTACATCTCGAACCATTTTATTTCAATGTGAAATCCCTGAACCTGGGA 721
 Db |||||
 Db 704 CATTGGAGCGCTACATCTCGAACCACTTTTATTTCAATGTGAAATCCCTGAACCTGGGA 762
 QY 722 TCAAGTCTATTACAGCTGTGGG 743
 Db |||||
 Db 763 TCAAGTCTATTACAGCTGTGGG 784

RESULT 7
 W00842
 LOCUS Yz96d12.rl Soares melanocyte 2NbHM Homo sapiens cDNA clone
 DEFINITION IMAGE:290903 5', mRNA sequence.
 ACCESSION W00842
 VERSION W00842.1 GI:1272822
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 371)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
 R., Williamson, A., Wohldmann, P., and Wilson, R.
 The WashU-Merck EST Project
 Unpublished (1995)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: ETPimer
 High quality sequence stop: 339.

Db 63 TTTTACTTCATTTTCTGAAGCTCTAGAGAAATTACAACCTTTGGTGATAACAAAA 10

/note="vector: pT7p3d-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NCI_CGAP_Sub3 library is a subtracted library derived from the NCI_CGAP_Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalised or subtracted NCI_CGAP libraries: NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10, NCI_CGAP_Co16, NCI_CGAP_Kid11, NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid5, NCI_CGAP_Kid2, NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_Co11, NCI_CGAP_Lu24, NCI_CGAP_Br23, NCI_CGAP_GC4, NCI_CGAP_GC6, NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6, NCI_CGAP_Br25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:

NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clones) 1322376-1323911, 1456008-1456775, 1500552-1502855; NCI_CGAP_Kid5 pool 1, LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clones) 1323912-1325831, 1471368-1472903, 1492104-1493255; NCI_CGAP_Lu5 pool 1, LLAM 3575-3582, 3851-3854 (IMAGE Clones) 1414920-1417991, 1520904-1522439; NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clones) 1257096-1258631, 1469064-1470983, 1475592-1476743; NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clones) 985608-986759, 1101192-1101959, 1217928-1220615; NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE Clones) 1057416-1061255, 1144584-1145351. Subtraction was performed as previously described (Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6, 791-806.

TAG_Lib=NCI_CGAP_Co10
TAG_Tissue=colon
TAG_SEQ=AAACG"

BASE COUNT 150 a 98 g 137 t
ORIGIN

Query Match 21.1%; Score 472.4; DB 10; Length 483;
Best Local Similarity 99.8%; Pred. No. 2.4e-80;
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1370 TCTCTGGAGACTCCACAGCTCTTTTGAATTCCTCAGTTTCCAGCTGGCTTTTATG 1429
DB 483 TCTCTGGAGACTCCACAGCTCTTTGAAATTCCTCAGTTTCCAGCTGGCTTTTATG 424

QY 1430 GGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTT 1489
DB 423 GGCTGGCAGTTGCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTT 364

QY 1490 TCAAGTGGCAGTTGCTCCAGCTTGTTCCTTCACATGCTCTTCATGTTGCTCC 1549
DB 363 TCAAGTGGCAGTTGCTCCAGCTTGTTCCTTCACATGCTCTTCATGTTGCTCC 304

QY 1550 TTTCCCTCTATTCCGACCCATTAGTACAGGATTGGCTTCGCTATCACTCTGACTGGAG 1609
DB 303 TTTCCCTCTATTCCGACCCATTAGTACAGGATTGGCTTCGCTATCACTCTGACTGGAG 244

QY 1610 TCCTCGGTATATCTCTTTATATGGACAAAGAACCCAGGTGTTTGAATAATGT 1669
DB 243 TCCTCGGTATATCTCTTTATATGGACAAAGAACCCAGGTGTTTGAATAATGT 184

QY 1670 CAGAGAAATTAACGAGCAATTAACAATAATCTGGAAGTTGTACGAGAAGATAAGT 1729
DB 183 CAGAGAAATTAACGAGCAATTAACAATAATCTGGAAGTTGTACGAGAAGATAAGT 124

QY 1730 TATGAACATAATGAGCTTGGCAATCTGCCACAGGGAGACACAAATAGGAT 1789
DB 123 TATGAACATAATGAGCTTGGCAATCTGCCACAGGGAGACACAAATAGGAT 64

QY 1790 TTTTACTTCATTTTCTGAAGCTCTAGAGAATTACAACCTTTGGTGATAACAAAA 1843

RESULT 5
BG326527 910 bp mRNA linear EST 27-FEB-2001
LOCUS 602425373f1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562994 5',
DEFINITION mRNA sequence.
ACCESSION BG326527 GI:13132964
VERSION BG326527.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue procurement: DCTD/DP
CDNA Library preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1276 row: b column: 19
High quality sequence stop: 706.
FEATURES
Location/Qualifiers
1..910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4562994"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 237 a 178 c 249 g 246 t
ORIGIN

Query Match 19.9%; Score 444.6; DB 12; Length 910;
Best Local Similarity 81.6%; Pred. No. 3.8e-75;
Matches 575; Conservative 0; Mismatches 74; Indels 56; Gaps 3;

QY 78 AAGAAGAAAAAGAGAGAGAGAGAAATTCAGGGCAATTTGGCATAGATTTTATCATAT 137
DB 63 AAGCAGAGGAAGACATCGATCAGTAACACACAGACACACCAAGTTGAAAGCTTTGTTTC 122

QY 138 TCTGGATTTTGGATTTCTTTTGTCTCTCATCTGAGTTCAGGAAGCCCTGTGTGTC 197
DB 123 TTTCCCTCTGTTTATTTTCCCGCTGTGCTCTACTATGTCAGAAAGCCTGTGTGTC 182

QY 198 CACCATCTCCAAAGAGAGTTTACCTGCGAGGAATCTTAACGGAGGCTGCCCTCCCTGGG 257
DB 183 CACCATCTCCAAAGAGAGTTTACCTGCGAGGAATCTTAACGGAGGCTGCCCTCCCTGGG 242

QY 258 CAACAAGGCCACCTGGCAGGAGCCCTTTTCAGGAAGAGACGCCTTTTTCAGGAAGAGA 317
DB 243 CAACAAGGCCACCTGGCAGGAGCCCTTTTCAGGAAGAGACGCCTTTTTCAGGAAGAGA 264

QY 318 CGCCTTTTTCAGGAAGAGAGAAAGTCAGCTGAAGAGAGAAAGTCACCTTTACTCAGGGAGT 377
DB 265 -----GAGAAAGTCAGCTGAAGAGAGAAAGTCACCTTTACTGAGGGAGT 308

mRNA sequence.
 BG490131
 VERSION BG490131.1 GI:13451641
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 919)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DP/Gapdar
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution by: Incyte Genomics, Inc.
 found through the I.M.A.G.E. Consortium information can be
 http://image.llnl.gov
 Plate: LCM1398 row: c column: 14
 High quality sequence stop: 726.
 Location/Qualifiers
 1..919
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4637509"
 /clone_lib="NIH_MGC_18"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 253 a 195 c 199 g 272 t
 ORIGIN

FEATURES
 source

Query Match 21.1%; Score 473; DB 12; Length 919;
 Best Local Similarity 76.2%; Pred. No. 1.4e-80;
 Matches 728; Conservative 0; Mismatches 5; Indels 222; Gaps 3;
 QY 829 GCAATCTGATATTAGTCCCTGGA-GTTATGAGCTTAATAAGGTCAACGAGAA 887
 DB 2 GCAATCTGATATTAGTCCCTGAGGTTATGAGCTTAATAAGGTCAACGAGAA 61
 QY 888 CTTTAAAGAGCGCTTTTTCAGGAAGAGATCAAGTATTACGGGTGCGACGTGCTTTTA 947
 DB 62 CTTTAAAGAGCGCTTTTTCAGGAAGAGATCAAGTATTACGGGTGCGACGTGCTTTTA 121
 QY 948 TTATGAATGTATGATGCTGGTGTGTTTACCTCAACTTTGTTACTGAAGAAGTAGA 1007
 DB 122 TTATGAATGTATGATGCTGGTGTGTTTACCTCAACTTTGTTACTGAAGAAGTAGA 181
 QY 1008 AAACCTGAAAAACCAUCCCTTGCATATGATATCAATGCCATTTGCACATTTG 1067
 DB 182 AAACCTGAAAA- 193
 QY 1068 CTATGCTGCAAAATGAGGCTACTTTACGACCATTAATGCTGAGGAGCTGCTTC 1127
 DB 194 - 193
 QY 1128 AAATGACGTGGCAGTGACCTTTTCTGAGCGGCTACTGGGAAATTCATAGCAGTTCC 1187
 DB 194 - 193
 QY 1188 GATCTTTGTCCTCTGCTGCTTTGGCTCCATGAAGGTTGGTGTGTTGCTGCTCCAG 1247
 DB 238 GATCTTTGTCCTCTGCTGCTTTGGCTCCATGAAGGTTGGTGTGTTGCTGCTCCAG 296

QY 1248 GTTATTTCTATGTTGCTCTCGAGAGGGTCACTTCCAGAAATCCTCTCCATGATTCATGT 1307
 DB 297 - 296
 QY 1308 CCACAAGCACACTCCCTTACCAGCTCTTATGTTTGGACCCCTTGGACAATGATAATGCT 1367
 DB 297 - 320
 QY 1368 CTTCTCTGAGAGCTCGACAGCTTTTGAATTTCTCAGTTTTCAGAGTGGCTTTTAT 1427
 DB 321 CTTCTCTGAGAGCTCGACAGCTTTTGAATTTCTCAGTTTTCAGAGTGGCTTTTAT 380
 QY 1428 TGGGCTGGCAGTTGCTGGGCTGATTATCTTCGATACAAATGCCAGATATGATCGTCC 1487
 DB 381 TGGGCTGGCAGTTGCTGGGCTGATTATCTTCGATACAAATGCCAGATATGATCGTCC 440
 QY 1488 TTTCAAGTGGCAGCTTTTCATCCAGCTTTTCTTCCATGATGCTCTTCTGATGCTG 1547
 DB 441 TTTCAAGTGGCAGCTTTTCATCCAGCTTTTCTTCCATGATGCTCTTCTGATGCTG 500
 QY 1548 CTTTCCCTCTTATTCGGACCCATTTAGTACAGGATTTGGCTTCATCATCTCTGACTGG 1607
 DB 501 CTTTCCCTCTTATTCGGACCCATTTAGTACAGGATTTGGCTTCATCATCTCTGACTGG 560
 QY 1608 AGTCCCTGCTTATTCCTTTTATTTATATATGACAGAAACCCAGGTGTTTGAATAAT 1667
 DB 561 AGTCCCTGCTTATTCCTTTTATTTATATATGACAGAAACCCAGGTGTTTGAATAAT 620
 QY 1668 GTACAGAAATATACAGACATTTACAAATTAATCTGGAAGTTGTACCAAGAGATATA 1727
 DB 621 GTACAGAAATATACAGACATTTACAAATTAATCTGGAAGTTGTACCAAGAGATATA 680
 QY 1728 GTTATGAATTAATGAGCTTTCAGATCTTGGCAATCTGCCCAAGGGGAGACACAAA 1782
 DB 681 GTTATGAATTAATGAGCTTTCAGATCTTGGCAATCTGCCCAAGGGGAGACACAAA 735

RESULT 4
 AW205572/c

LOCUS
 DEFINITION

ACCESSION
 VERSION

KEYWORDS
 SOURCE

ORGANISM

REFERENCE

AUTHORS
 TITLE

JOURNAL
 COMMENT

483 bp mRNA linear EST 02-DEC-1999
 UI-H-B11-adt-h-02-0-UI-sl NCI-CGAP-Sub3 Homo sapiens cDNA clone
 IMAGE:2718171 3', mRNA sequence.

EST.
 GI:6505046

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 483)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:

NCI-CGAP clone distribution information can be found through the

www-bio.llnl.gov/bbrp/image/image.html

Seg primer: M13 Forward

POLYA-Yes.

location/Qualifiers
 1..483

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2718171"

/clone_lib="NCI-CGAP-Sub3"

/lab_host="DH10B (Life Technologies)"

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 05:48:49 ; Search time 3548 Seconds
(without alignments)
10220.320 Million cell updates/sec

Title: US-09-667-170a-440

Perfect score: 2239

Sequence: 1 ggagggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_estl: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pin: *
21: em_gss_vit: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_fod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	512.8	22.9	516	9	AL702210
c 2	508.4	22.7	511	14	BM887840
3	473	21.1	919	12	BM887840 TMT023 Hu
c 4	472.4	21.1	483	10	BG490131
5	444.6	19.9	910	12	AW205572
6	411.2	18.4	936	12	BG326527
					BG284503
					BG284503 602425373
					BG284503 602408645

7	363	16.2	371	14	W00842
8	359.8	16.1	497	14	BQ370588
9	326.2	14.6	1072	12	BG388107
10	311.6	13.9	1962	11	BC022457
11	250	11.2	403	13	BQ075046
12	243.2	10.9	872	14	BQ430089
13	228.4	10.2	945	14	BQ682698
14	225.4	10.1	524	12	BF286052
15	221.8	9.9	912	14	BQ684656
16	221.6	9.9	924	14	BQ894656
17	220.2	9.8	1182	11	AK005282
18	213.2	9.5	1068	14	BM920139
19	212.4	9.5	915	14	BQ211976
20	207	9.2	952	14	BQ680741
c 21	204.2	9.1	557	17	AQ002318
22	202	9.0	959	9	AL553189
23	201.8	9.0	876	14	BQ677699
24	201	9.0	1123	13	BM460946
25	200	8.9	682	9	AJ392529
26	200	8.9	702	9	AL635772
27	198.2	8.9	831	13	BI181265
28	197.2	8.8	991	13	BM556677
29	193.4	8.6	945	14	BQ682351
30	191.8	8.6	940	14	BQ678193
31	189.4	8.5	915	14	BQ681717
32	187.2	8.4	649	9	AL635943
c 33	186	8.3	530	9	AL102370
34	184	8.2	637	9	AL586948
35	182.8	8.2	581	12	BF717670
36	181.6	8.1	1155	13	BM474488
37	180.6	8.1	889	9	AL522431
38	180	8.0	949	14	BQ678681
39	179.8	8.0	927	14	BQ683707
40	179.4	8.0	1054	14	BM919004
41	177.4	7.9	706	12	BG761990
42	173.4	7.7	953	14	BQ683115
43	172.8	7.7	952	9	AL560263
44	172	7.7	551	14	BM962577
45	171.8	7.7	701	14	BM963289

ALIGNMENTS

RESULT 1
AL702210
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL702210 516 bp mRNA linear EST 22-MAR-2002
DKFZp686C09156_r1 686 (synonym: hlcc3) Homo sapiens CDNA clone
AL702210
EST
GI:19685565

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 516)

Bloecker, H., Boecker, M., Brandt, P., Mewes, W., Weil, B. and Wiemann, S.

EST (Bloecker, H., Boecker, M., Brandt, P., Mewes, W., Weil, B. and Wiemann, S.)

Unpublished (1999)

Contact: Bloecker H

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

Sequenced by GBF (National Research Centre for Biotechnology Ltd.,

Braunschweig/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No sl sequence available.

This clone (DKFZp686C09156) is available at the RZPD in Berlin.

Search completed: April 16, 2003, 11:21:56
Job time : 528 secs

PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249266.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 XX
 DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 XX useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PT
 XX
 PS Disclosure; SEQ ID NO 23451; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to prevent,
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 123 BP; 32 A; 31 C; 37 G; 23 T; 0 other;
 Query Match 4.5%; Score 100; DB 22; Length 123;
 Best Local Similarity 100.0%; Pred. No. 7.3e-32;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 182 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGTTACTCGAGGAAATGTTACCGGA 241
 Db 14 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGTTACTCGAGGAAATGTTACCGGA 73
 QY 242 GGCTGCCCTTCCTGGGCAACAGGAGCCACCTGGGCGAGGA 281
 Db 74 GGCTGCCCTTCCTGGGCAACAGGAGCCACCTGGGCGAGGA 113
 RESULT 15
 ABN35211
 ID ABN35211 standard; DNA; 60 BP.
 XX
 AC ABN35211;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human spliced transcript detection oligonucleotide SEQ ID NO:7959.
 XX Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX Homo sapiens.
 OS WO200210449-A2.
 XX
 PN 07-FEB-2002.
 PD
 XX 20-JUL-2001; 2001WO-IB01903.
 PF
 XX 28-JUL-2000; 2000US-221607P.
 PR

527 TGGGAACAACTATATAAGAAATCTGGAGGTCATTACACATATATTTGG 574
|||||
636 TGGGAACAACTATATAAGAAATCTGGAGGTCATTACACATATATTTGG 683
|||||

SULT 13
L37718
ABL37718 standard; cDNA; 575 BP.

ABL37718; 08-APR-2002 (first entry)
Human colon tumour antigen polynucleotide SEQ ID NO:1307.
Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
colon tumour metastatic antigen; diagnosis; gene; ss.
Homo sapiens.
WO200196388-A2.
20-DEC-2001.
08-JUN-2001; 2001WO-US18557.
09-JUN-2000; 2000US-210899P.
20-FEB-2001; 2001US-270216P.
(CORI-) CORIXA CORP.
Jiang Y, Harlocker SL, Secríst H;
WPI; 2002-114514/15.
Novel isolated colon tumor polynucleotide differentially expressed in
colon tumor or colon metastatic tumor and polypeptides encoded by them,
useful for inhibiting development of cancer in patient -
Claim 1; SEQ ID 1307; 105pp; English.
ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
which were isolated from human colon tumour and colon metastatic tumour
cDNA libraries. (I) have cytostatic activity and can be used in vaccine
production. (I) can be used for stimulating and/or expanding T cells
specific for a tumour protein on contact with the T cells. They are also
useful for inhibiting the development of cancer in a patient. (I) can be
used as probes or primers for nucleic acid hybridisation, for preparing
mutant species primers, or primers for use in genetic constructions. (I)
can be used in the diagnosis of a colon tumour.
Sequence 575 BP; 151 A; 122 C; 153 G; 141 T; 8 other;
Query Match 4.8%; Score 107; DB 24; Length 575;
Best Local Similarity 100.0%; Pred. No. 7.9e-35;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
334 GAGAAAGTGCAGCTGAAGAGGAAAGTCATTACTGAGGGAGCTCCATTATCATTTGGC-393
|||||
354 GAGAAAGTGCAGCTGAAGAGGAAAGTCATTACTGAGGGAGCTCCATTATCATTTGGC 413
|||||
394 ACCATCATTTGAGCAGGAAATCTTCATCTCCCTAAGGCGTCTCCA 440
|||||
414 ACCATCATTTGAGCAGGAAATCTTCATCTCCCTAAGGCGTCTCCA 460
|||||

RESULT 14
AAK68639
ID AAK68639 standard; DNA; 123 BP.
XX
AC AAK68639;
XX

XX PD 20-DEC-2001.
 XX PF 08-JUN-2001; 2001WO-US18577.
 XX PR 09-JUN-2000; 2000US-210821P.
 XX PR 18-DEC-2000; 2000US-256571P.
 XX PR 10-MAY-2001; 2001US-290240P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;
 XX PT WPI; 2002-139708/18.
 XX PT Novel isolated polynucleotide encoding a polypeptide comprising a
 PT portion of colon tumour protein, useful for detection, diagnosis and
 PT therapy of human colon cancer.
 XX PS Claim 1; Page 174; 220pp; English.
 XX CC The invention relates to an isolated polynucleotide (I) encoding a
 CC polypeptide (II) comprising at least a portion of a colon tumour
 CC protein. (I), (II) and antibody (III) to (II) are useful for determining
 CC the presence of a cancer in a patient. (I), (II) or antigen presenting
 CC cells expressing (I) is useful for stimulating and/or expanding T cells
 CC specific for a tumour protein, by contacting T cells with (I), (II) or
 CC antigen-presenting cells that express (I), under conditions and for a
 CC time sufficient to permit the stimulation and/or expansion of T cells.
 CC (I), (II), or antigen presenting cells that express (II) are useful for
 CC treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
 CC isolated from a patient with (I), (II) or antigen presenting cells that
 CC express (II), such that T cells proliferate, and administering to the
 CC patient an effective amount of the proliferated T cells, thus inhibiting
 CC the development of a cancer in the patient. (I) or (II) is useful in
 CC vaccines and pharmaceutical compositions for prevention and treatment
 CC of colon malignancies and for the diagnosis and monitoring of such
 CC cancers. (I), (II) or (III) is useful for detection, diagnosis and/or
 CC therapy of human colon cancer. (I) is useful as a probe or primer for
 CC nucleic acid hybridisation, and in the design and preparation of
 CC ribozyme molecules for inhibiting expression of (II) in tumour cells.
 CC ABK27564-ABK27807 represent novel human colon cancer coding
 CC sequences and primers of the invention.
 XX SQ Sequence 596 BP; 160 A; 151 C; 126 G; 158 T; 1 other;
 Query Match 12.0%; Score 269; DB 24; Length 596;
 Best Local Similarity 100.0%; Pred. No. 1.5e-102;
 Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 334 GAGAAAGTCAGCTGAAGAGAAAGTCATTTACTGAGGGAGTCTCCATTATCATTTGGC 393
 DB 269 GAGAAAGTCAGCTGAAGAGAAAGTCATTTACTGAGGGAGTCTCCATTATCATTTGGC 210
 QY 394 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGCGGTGCTCCAGAACCGGGCAGC 453
 DB 209 ACCATCATTTGGAGCAGGAATCTTCATCTCTCTAAGGCGGTGCTCCAGAACCGGGCAGC 150
 QY 454 GTGGGCATCTCTGACCATCTGGAGCGTGTGGGTCTCTGACATTTTGGAGCTTTG 513
 DB 149 GTGGGCATCTCTGACCATCTGGAGCGTGTGGGTCTCTGACATTTTGGAGCTTTG 90
 QY 514 TCTTATGCTGAATTTGGGAACACTATAAAGAAATCTGGAGTCAATACACATATATTTTG 573
 DB 89 TCTTATGCTGAATTTGGGAACACTATAAAGAAATCTGGAGTCAATACACATATATTTTG 30
 QY 574 GAAGTCTTTGGTCCATTACAGCTTTTCT 602
 DB 29 GAAGTCTTTGGTCCATTACAGCTTTTCT 1
 RESULT 12
 AAZ16609

ID AAZ16609 standard; cDNA; 800 BP.
 XX AC AAZ16609;
 XX DT 12-OCT-1999 (first entry)
 XX DE Human gene expression product cDNA sequence SEQ ID NO:4079.
 XX KW Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX OS Homo sapiens.
 XX PN WO9938972-A2.
 XX PD 05-AUG-1999.
 XX PF 28-JAN-1999; 99WO-US01619.
 XX PR 03-APR-1998; 98US-0080666.
 PR 24-JAN-1998; 98US-0072910.
 PR 24-FEB-1998; 98US-0075954.
 PR 31-MAR-1998; 98US-0080114.
 XX 03-APR-1998; 98US-0080515.
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 PI CrKvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lamson G, Leskowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 XX WPI; 1999-494092/41.
 XX PT Novel human genes and their expression products which are
 PT differentially expressed in different cell types
 XX Claim 1; Page 1934; 2479pp; English.
 CC The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
 CC polynucleotides can be used as a source of primers and probes, which can
 CC be used for a variety of purposes, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensics, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and
 CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels between
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
 CC identify a genetic predisposition or susceptibility to a disease such as
 CC cancer). The polynucleotides of the invention are especially used in the
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
 CC and lung cancer. The polynucleotides can also be used to screen for
 CC peptide analogues and antagonists.
 XX SQ Sequence 800 BP; 199 A; 166 C; 202 G; 211 T; 22 other;
 Query Match 4.8%; Score 108; DB 20; Length 800;
 Best Local Similarity 100.0%; Pred. No. 2.9e-35;
 Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 467 TGACCATCTGACGCTGTGGGGTCTCTGACATTTTGGAGCTTTGCTTATGCTGAAT 526
 DB 576 TGACCATCTGACGCTGTGGGGTCTCTGACATTTTGGAGCTTTGCTTATGCTGAAT 635

Thu Apr 17 09:48:28 2003

```

XX DR AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
XX CC libraries constructed from human colon cancer cell lines. The present
XX CC invention also describes a method of detecting differentially expressed
XX CC genes correlated with a cancerous state of a mammalian cell, comprising
XX CC detecting at least one differentially expressed gene product in a test
XX CC sample derived from a cell suspected of being cancerous, where detection
XX CC of the differentially expressed gene product is correlated with a
XX CC cancerous state of the cell from which the test sample was derived.
XX CC The polynucleotide sequences can be used in a method for detecting
XX CC differentially expressed genes correlated with a cancerous state of a
XX CC mammalian cell. The polynucleotides can also be used as probes for
XX CC detecting and mapping related genes. They can be used in diagnosis and
XX CC prognosis of diseases and disorders (e.g. identification of
XX CC pre-metastatic or metastatic cancerous states, stages of cancer, or
XX CC responsiveness of cancer to therapy). This is particularly for breast
XX CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
XX CC negative breast cancer, lung cancer, and colon cancer.
XX CC
XX CC Sequence 300 BP; 80 A; 65 C; 66 G; 89 T; 0 other;
XX SQ

```

Query Match 12.7%; Score 284; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 8.6e-109;
 Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 658 CTGGCATTGGACGCTACATCTCGAACCAATTTTATTTCAATGTGAATCCCTGAACCT 717
DB 17 CTGGCATTGGACGCTACATCTCGAACCAATTTTATTTCAATGTGAATCCCTGAACCT 76
QY 718 GCGATCAGCTCATTACAGCTGTGGCATACTGTAGTGATGGTCTCTAAATAGCATGAGT 777
DB 77 GCGATCAGCTCATTACAGCTGTGGCATACTGTAGTGATGGTCTCTAAATAGCATGAGT 136
QY 778 GTCAGCTGGAGCGCCCGGATCCAGATTTTCACTTTTCAAGCTTCACAGCAATTCG 837
DB 137 GTCAGCTGGAGCGCCCGGATCCAGATTTTCACTTTTCAAGCTTCACAGCAATTCG 196
QY 838 ATAAATATAGTCCCTGGAGTATGACGCTAATTAAGGTCAACGCGAGAACTTTAAAGAC 897
DB 197 ATAAATATAGTCCCTGGAGTATGACGCTAATTAAGGTCAACGCGAGAACTTTAAAGAC 256
QY 898 GCCTTTTCAGGAAGAGATTCAAGTATTACGGGGTTGCCACTGGC 941
DB 257 GCCTTTTCAGGAAGAGATTCAAGTATTACGGGGTTGCCACTGGC 300

```

RESULT 10
 AAD23462
 ID AAD23462 standard; cDNA; 337 BP.
 XX AC AAD23462;
 XX DT 26-FEB-2002 (first entry)
 XX DE Human lung tumour-specific 20E10 5' cDNA.
 XX KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
 XX KW antisense-therapy; vaccine; immune response; lung cancer; 20E10; ss.
 XX OS Homo sapiens.
 XX PN W0200172295-A2.
 XX PD 04-OCT-2001.
 XX PF 28-MAR-2001; 2001WO-US09991.

```

XX PR 29-MAR-2000; 2000US-0538037.
XX PR 05-JUN-2000; 2000US-0588937.
XX PR 18-AUG-2000; 2000US-0640878.
XX PR 22-SEP-2000; 2000US-234517P.
XX PR 01-NOV-2000; 2000US-0704512.
XX PR 14-DEC-2000; 2000US-0738973.
XX PA (CORI-) CORIXA CORP.
XX PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CV;
XX PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX PI WPI; 2001-639201/73.
XX DR New human lung-specific polynucleotides and polypeptides for the
XX DR diagnosis and treatment of disease e.g. lung cancer -
XX PT Claim 1; Page 334; 378pp; English.
XX PS The invention relates to isolated lung tumour-specific proteins and
XX PS their corresponding cDNA molecules. Lung tumour-specific proteins and
XX PS their antigen-presenting cells are useful for stimulating and/or
XX PS expanding T cells specific for a tumour protein, and for inhibiting
XX PS the development of cancer. The invention also relates to a composition
XX PS useful for stimulating an immune response, and for treating cancer. The
XX PS lung tumour specific oligonucleotide is useful in gene therapy and for
XX PS diagnosis, detection and treatment of lung cancer. The present sequence
XX PS is a cDNA encoding human lung tumour-specific protein.
XX SQ Sequence 337 BP; 103 A; 60 C; 93 G; 81 T; 0 other;
XX SQ

```

Query Match 12.6%; Score 281; DB 23; Length 337;
 Best Local Similarity 100.0%; Pred. No. 1.5e-107;
 Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGAGTTGAAGTGAAGAGAGATCATGCCAGCTGGGTGACAGTGAGACTCTGTCTCAAAAC 60
DB 4 GGAGTTGAAGTGAAGAGAGATCATGCCAGCTGGGTGACAGTGAGACTCTGTCTCAAAAC 63
QY 61 AGAATTAAGGAAAGAAAGAAAGAAAGAGAGAGAGAGAAATTCAGGCCAATTTGG 120
DB 64 AGAATTAAGGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAAATTCAGGCCAATTTGG 123
QY 121 CATACATTTTATCATATTTCTGGATTTTGGATTTTCTTCTCATCTGATTCATCA 180
DB 124 CATACATTTTATCATATTTCTGGATTTTGGATTTTCTTCTCATCTGATTCATCA 183
QY 181 GGAAAGCCCTGTTGTGTCCACCATCTCCAAAGAGAGGTTACCTGCAGGAAATTTAACGGG 240
DB 184 GGAAAGCCCTGTTGTGTCCACCATCTCCAAAGAGAGGTTACCTGCAGGAAATTTAACGGG 243
QY 241 AGGCTGCCCTTCCCTGGGCAACAGGAGCCACCTGGGCAGGA 281
DB 244 AGGCTGCCCTTCCCTGGGCAACAGGAGCCACCTGGGCAGGA 284

```

RESULT 11
 ABR27651/C
 ID ABR27651 standard; cDNA; 596 BP.
 XX AC ABR27651;
 XX DT 09-APR-2002 (first entry)
 XX DE Human colon cancer expressed sequence tag, Seq ID no 88.
 XX KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
 XX KW expressed sequence tag.
 XX OS Homo sapiens.
 XX PN W0200196390-A2.

QY 899 CCTTTTCCAGGAGAGATTCAAGTATTACGGGTGCGCACTGGCTTTTATTATGGAATGT 958
 Db |||||||
 QY 61 CCTTTTCCAGGAGAGATTCAAGTATTACGGGTGCGCACTGGCTTTTATTATGGAATGT 120
 Db |||||||
 QY 959 ATGCATATGCTGGCTGGTGTTCACCTCAACTTGTACTGAAGAGTAGAAAACCCCTGAAA 1018
 Db |||||||
 QY 121 ATGCATATGCTGGCTGGTGTTCACCTCAACTTGTACTGAAGAGTAGAAAACCCCTGAAA 180
 Db |||||||
 QY 1019 AAACCAATCCCTTGCATATGATATATCCATGCGCATGTCACCAATGGCTATGCTGCTGA 1078
 Db |||||||
 QY 181 AAACCAATCCCTTGCATATGATATATCCATGCGCATGTCACCAATGGCTATGCTGCTGA 240
 Db |||||||
 QY 1079 CAATATGGCTTACCTACGACCAATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGG 1138
 Db |||||||
 QY 241 CAATATGGCTTACCTTACGACCAATTAATGCTGAGGAGCTGCTGCTTTCAAAATGCAGTGG 300
 Db |||||||
 RESULT 8
 ID AAZ14304
 AC AAZ14304 standard; cDNA; 288 BP.
 XX AAZ14304;
 DT 12-OCT-1999 (first entry)
 DE Human gene expression product cDNA sequence SEQ ID NO:1773.
 KW Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 OS Homo sapiens.
 PN W09938972-A2.
 PD 05-AUG-1999.
 XX 28-JAN-1999; 99WO-US01619.
 XX 03-APR-1998; 98US-0080666.
 PR 28-JAN-1998; 98US-0072910.
 PR 24-FEB-1998; 98US-0075954.
 PR 31-MAR-1998; 98US-0080114.
 PR 03-APR-1998; 98US-0080515.
 XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 XX WPI; 1999-494092/41.
 XX Novel human genes and their expression products which are
 PT differentially expressed in different cell types
 PS Claim 1; Page 1061; 2479pp; English.

CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels between
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
 CC identify a genetic predisposition or susceptibility to a disease such as
 CC cancer). The polynucleotides of the invention are especially used in the
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
 CC and lung cancer. The polynucleotides can also be used to screen for
 CC peptide analogues and antagonists.

SQ Sequence 288 BP; 81 A; 58 C; 59 G; 90 T; 0 other;

Query Match 12.9%; Score 288; DB 20; Length 288;

Best Local Similarity 100.0%; Pred. No. 1.8e-110;

Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 839 TAATTATAGTCCCTGGAGTTATGCGCTAATTAAGGTCAAACGAGACATTTAAAGACG 898
 Db |||||||
 QY 899 CTTTTTCAGGAGAGATTCAAGTATTACGCGGTGCGCACTGGCTTTTATTATGGAATGT 958
 Db |||||||
 QY 61 CTTTTTCAGGAGAGATTCAAGTATTACGCGGTGCGCACTGGCTTTTATTATGGAATGT 120
 Db |||||||
 QY 959 ATGCATATGCTGGCTGGTGTTCACCTCAACTTGTACTGAAGAGTAGAAAACCCCTGAAA 1018
 Db |||||||
 QY 1019 AAACCAATCCCTTGCATATGATATATCCATGCGCATGTCACCAATGGCTATGCTGCTGA 1078
 Db |||||||
 QY 181 AAACCAATCCCTTGCATATGATATATCCATGCGCATGTCACCAATGGCTATGCTGCTGA 240
 Db |||||||
 QY 1079 CAATATGGCTTACCTACGACCAATTAATGCTGAGGAGCTGCTGCTTT 1126
 Db |||||||
 QY 241 CAATATGGCTTACCTTACGACCAATTAATGCTGAGGAGCTGCTGCTTT 288

RESULT 9

ID AAZ00880

XX AAZ00880 standard; cDNA; 300 BP.

AC AAZ00880;

XX 19-MAY-2000 (first entry)

DT Human colon cancer cell line polynucleotide sequence SEQ ID NO:871.

DE Human; colon cancer; tumour; diagnosis; gene expression product;

KW probe; detection; cancerous state; metastasis; identification;

KW breast cancer; oestrogen receptor-positive breast cancer; therapy;

KW oestrogen receptor-negative breast cancer; lung cancer; ss.

OS Homo sapiens.

XX W09958675-A2.

XX 18-NOV-1999.

XX 13-MAY-1999; 99WO-US10602.

XX 14-MAY-1998; 98US-0085426.

PR 15-MAY-1998; 98US-0085537.

PR 15-MAY-1998; 98US-0085696.

PR 21-OCT-1998; 98US-0105234.

PR 27-OCT-1998; 98US-0105877.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;

PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;

PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;

CC The present invention describes a library of human polynucleotides

CC comprising the sequences given in AAZ12532 to AAZ1779. Also described is

CC a method of detecting differentially expressed genes correlated with the

CC cancerous state of a mammalian cell, comprising detecting at least one

CC differentially expressed gene product in a test sample from a cell

CC suspected of being cancerous, where the gene product is encoded by one

CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ1779. The

CC polynucleotides can be used as a source of primers and probes, which can

CC be used for a variety of purpose, e.g. detection of expression levels,

CC mapping, tissue typing or profiling, forensics, genetic analysis and

CC detection of polymorphisms. Polypeptides encoded by the polynucleotides

CC can be used for raising antibodies for experimental, diagnostic and

[illegible]

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition; ss.
 OS Homo sapiens.
 XX WO200154477-A2.
 PN 02-AUG-2001.
 XX 25-JAN-2001; 2001WO-US02687.
 XX 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 XX WPI; 2001-476164/51.
 DR P-PSDB; AAM23914.
 XX Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -
 XX Claim 1; Page 484; 1275pp; English.
 XX The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA
 CC of the invention.
 XX Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;
 XX
 Query Match 76.8%; Score 1719; DB 22; Length 1958;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1769; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 334 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGAGTCTCCATTATCATTTGGC 393
 DB 168 GAGAAAGTGCAGCTGAAGAGGAAAGTCACTTACTGAGGGAGTCTCCATTATCATTTGGC 227
 QY 394 ACCATCATTGGAGCAGGAATCTTCTCTAAGGGCGTGTCCAGAACACGGGCGAGC 453
 DB 228 ACCATCATTGGAGCAGGAATCTTCTCTAAGGGCGTGTCCAGAACACGGGCGAGC 287
 QY 454 GTGGGCATGCTCTGACCATCTGGAGGCTGTGTGGGGTCTGTCTCACTATTTGGAGCTTG 513
 DB 288 GTGGGCATGCTCTGACCATCTGGAGGCTGTGTGGGGTCTGTCTCACTATTTGGAGCTTG 347
 QY 514 TCTTATGCTGAATGGGAACAACTATAAAGAAATCTGGAGTCAATACACATATTTTG 573
 DB 348 TCTTATGCTGAATGGGAACAACTATAAAGAAATCTGGAGTCAATACACATATTTTG 407
 QY 574 GAAGTCTTGGTCCATTACCACTTTTGTACGAGTCTGGGTGGAACCTCTCATATATACGC 633
 DB 408 GAAGTCTTGGTCCATTACCACTTTTGTACGAGTCTGGGTGGAACCTCTCATATATACGC 467
 QY 634 CTTGCAGCTACTGCTGTGATATCCCTGGCATTTGGAGCTACATCTCGAACCATTTT 693
 DB 468 CTTGCAGCTACTGCTGTGATATCCCTGGCATTTGGAGCTACATCTCGAACCATTTT 527
 QY 694 ATTCATGTGAATCCCTGNACTTGGATCAAGCTCATTACAGCTGTGGGCATCACTGTA 753

DB 528 ATTCAATGTGAATCCCTGAACCTTGGATCAAGCTCATTACAGCTGTGGGCATCACTGTA 587
 QY 754 GTGATGTCTTAATAGCATGAGTGTGAGCGCCGGATCCAGATTTCTTAACC 813
 DB 588 GTGATGTCTTAATAGCATGAGTGTGAGCGCCGGATCCAGATTTCTTAACC 647
 QY 814 TTTTGCAGGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGAGCTAATAA 873
 DB 648 TTTTGCAGGCTCACAGCAATTTCTGATAATATAGTCCCTGGAGTTATGAGCTAATAA 707
 QY 874 GGTCAACGCGAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTACGCGTTG 933
 DB 708 GGTCAACGCGAGAACTTTAAAGACGCTTTTCAGGAAGAGATTCAAGTATTACGCGTTG 767
 QY 934 CCACCTGGCTTTTATATGGAATGATGATATGCTGGCTGGTTTACCTCAACTTTGTT 993
 DB 768 CCACCTGGCTTTTATATGGAATGATGATATGCTGGCTGGTTTACCTCAACTTTGTT 827
 QY 994 ACTGAAGAAGTAGAAACCCCTGAAACACCAATCCCTTGCATATATATCATCGGCC 1053
 DB 828 ACTGAAGAAGTAGAAACCCCTGAAACACCAATCCCTTGCATATATATCATCGGCC 887
 QY 1054 ATTGTCAACCATTTGCTATGCTGACAAATGTCCTACTTTAGACCATTAATGCTGAG 1113
 DB 888 ATTGTCAACCATTTGCTATGCTGACAAATGTCCTACTTTAGACCATTAATGCTGAG 947
 QY 1114 GAGCTGCTGCTTTCAAAATGAGTGGCAGTGGCAGCTTTCTGAGCGGCTACTGGGAAATTC 1173
 DB 948 GAGCTGCTGCTTTCAAAATGAGTGGCAGTGGCAGCTTTCTGAGCGGCTACTGGGAAATTC 1007
 QY 1174 TCATTAGCAGTTCGCGATCTTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1233
 DB 1008 TCATTAGCAGTTCGCGATCTTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1067
 QY 1234 TTTGCTGCTCTCCAGGTTATTTCTATGTTGGCTTCGAGAGGTCACCTTCCAGAAATCTC 1293
 DB 1068 TTTGCTGCTCTCCAGGTTATTTCTATGTTGGCTTCGAGAGGTCACCTTCCAGAAATCTC 1127
 QY 1294 TCATGATTCATCTCCGAGCAGTCTCTACACAGCTGTTATTTGTTTGCACCTTTG 1353
 DB 1128 TCCATGATTCATCTCCGAGCAGTCTCTACACAGCTGTTATTTGTTTGCACCTTTG 1187
 QY 1354 ACAATGATAATGCTCTCTCGAGAGCTCGAGAGTCTTTTGAATTTCTCAGTTTGGC 1413
 DB 1188 ACAATGATAATGCTCTCTCGAGAGCTCGAGAGTCTTTTGAATTTCTCAGTTTGGC 1247
 QY 1414 AGTGGCTTTTATTTGGCTGGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
 DB 1248 AGTGGCTTTTATTTGGCTGGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1307
 QY 1474 GATATGATCTCTCTCAAGGTGCCACTGTTCATCCAGCTTTGTTTCTTCCATCATGC 1533
 DB 1308 GATATGATCTCTCTCAAGGTGCCACTGTTCATCCAGCTTTGTTTCTTCCATCATGC 1367
 QY 1534 CTCTTCTATGCTGCTTCTCTATTCGAGCCCTTATAGTACAGGATTTGCTTCTGCTG 1593
 DB 1368 CTCTTCTATGCTGCTTCTCTATTCGAGCCCTTATAGTACAGGATTTGCTTCTGCTG 1427
 QY 1594 ATCACTCTGAGTGGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1653
 DB 1428 ATCACTCTGAGTGGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1487
 QY 1654 TGGTTTAGAATTAATGTCAGAGAAATAACAGAACATTAACAAATTAATCTGGAAGTTGTA 1713
 DB 1488 TGGTTTAGAATTAATGTCAGAGAAATAACAGAACATTAACAAATTAATCTGGAAGTTGTC 1547
 QY 1714 CCAGAGAAAGATAGTTATGAACCTAGTCTGAGATCTTGGCAATCTGCCCAAGGGGA 1773
 DB 1548 CCAGAGAAAGATAGTTATGAACCTAGTCTGAGATCTTGGCAATCTGCCCAAGGGGA 1607
 QY 1774 GACACAAAATAGGGAATTTTACTTCTTCTGAAAGTCTAGAGAAATTAACAATTTGGTG 1833
 DB 1608 GACACAAAATAGGGAATTTTACTTCTTCTGAAAGTCTAGAGAAATTAACAATTTGGTG 1667

bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention.

SQ Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;

```

Query Match          76.8%; Score 1719; DB 22; Length 1958;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1769; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY	334	GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTAGGGGAGTCTCCATTATCATTTGCC	393
Db	168	GAGAAAGTGCAGCTGAAGAGAAAGTCACTTTACTAGGGGAGTCTCCATTATCATTTGCC	227
QY	394	ACCATCATTTGGAGCAGGAATCTTCATCTCTCCCTTAAGGGCGTGTCCAGAACACGGGCAGC	453
Db	228	ACCATCATTTGGAGCAGGAATCTTCATCTCTCCCTTAAGGGCGTGTCCAGAACACGGGCAGC	287
QY	454	GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCACATTTTGGAGCTTTG	513
Db	288	GTGGGCATGTCTCTGACCATCTGGACGGTGTGTGGGGTCTCTGTCACATTTTGGAGCTTTG	347
QY	514	TCTTATGCTGAATTTGGGAAACAACATATAAGAAATCTGGAGGTCATTACACATATATTTTG	573
Db	348	TCTTATGCTGAATTTGGGAAACAACATATAAGAAATCTGGAGGTCATTACACATATATTTTG	407
QY	574	GAACTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGTGGAACCTCTCATATATATTTT	633
Db	408	GAACTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGTGGAACCTCTCATATATATTTT	467
QY	634	CCTCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTCTGGAACCATTTTTT	693
Db	468	CCTCAGCTACTGCTGTGATATCCCTGGCATTTGGACGCTACATCTCTGGAACCATTTTTT	527
QY	694	ATTCAATGTGAAATCCCTGAACTTGGCGATTCGAGTCATTACAGCTGTGGGCGATTAAC	753
Db	528	ATTCAATGTGAAATCCCTGAACTTGGCGATTCGAGTCATTACAGCTGTGGGCGATTAAC	587
QY	754	GTGATGGTCTTAATAGCATGAGTGTACGCTGGAGCGCCGGATCCAGATTTTCTTAACC	813
Db	588	GTGATGGTCTTAATAGCATGAGTGTACGCTGGAGCGCCGGATCCAGATTTTCTTAACC	647
QY	814	TTTTGCAAGCTCACGCAATTTCTGATAATATATAGTCCCTGGAGTATGCAAGCTTAAT	873
Db	648	TTTTGCAAGCTCACGCAATTTCTGATAATATATAGTCCCTGGAGTATGCAAGCTTAAT	707
QY	874	GGTCAACCCGAACTTTAAGACGCCCTTTTCAGGAAGAGATTCAAGTATTACCGGTTG	933
Db	708	GGTCAACCCGAACTTTAAGAGGCCCTTTTCAGGAAGAGATTCAAGTATTACCGGTTG	767
QY	934	CCACTGGCTTTTATTATGGAAATGTATGCAATATCTGGCTGGTTTACCTCAACTTTGTT	993
Db	768	CCACTGGCTTTTATTATGGAAATGTATGCAATATCTGGCTGGTTTACCTCAACTTTGTT	827
QY	994	ACTGAAGAGTAGAAAACCCCTGAAAACCATTTCCCTTGCAATATGTATATCCATGGCC	1053
Db	828	ACTGAAGAGTAGAAAACCCCTGAAAACCATTTCCCTTGCAATATGTATATCCATGGCC	887
QY	1054	ATTGTACCAATTTGGCTATGTCTGACAAATGTGGCTACTTTTCTGAGCGGCTACTGGGAAATTC	1113
Db	888	ATTGTACCAATTTGGCTATGTCTGACAAATGTGGCTACTTTTCTGAGCGGCTACTGGGAAATTC	947
QY	1114	GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGAACCTTTTCTGAGCGGCTACTGGGAAATTC	1173
Db	948	GAGCTGCTGCTTTCAAAATGCAGTGGCAGTGAACCTTTTCTGAGCGGCTACTGGGAAATTC	1007
QY	1174	TCATTAGCAGTTCCGATCTTTGTTGCCCTCTCTCTGCTTTGGCTCCATGAACGGTGGTGTG	1233

RESULT 4	AAH98573	AAH	AAH	12-	Hum
ID	XX	AC	XX	DT	XX
DE					

12-OCT-2001 (first entry)

Human EST-derived coding sequence SEQ ID NO: 430.

QY 1234 TTTCGCTCTCCAGGTTATCTATGTTGGTCTCGAGAGGGTCACCTTCCAGAAATCCTC 1293
 DB 1178 TTTCGCTCTCCAGGTTATCTATGTTGGTCTCGAGAGGGTCACCTTCCAGAAATCCTC 1237
 QY 1294 TCCATGATTCATGTCGCGAAGCACACTCCTTACCAAGCTGTTTATGTTTGGACCCCTTG 1353
 DB 1238 TCCATGATTCATGTCGCGAAGCACACTCCTTACCAAGCTGTTTATGTTTGGACCCCTTG 1297
 QY 1354 ACAATGATATGCTCTCTCGAGACCTCGACAGCTCTTTGAATTTCTCAGTTTGGCC 1413
 DB 1298 ACAATGATATGCTCTCTCGAGACCTCGACAGCTCTTTGAATTTCTCAGTTTGGCC 1357
 QY 1414 AGGTGGCTTTTATGTTGGCTGGCAGTGTGCTGCTGATTTATCTTCGATACAAATGCCCA 1473
 DB 1358 AGGTGGCTTTTATGTTGGCTGGCAGTGTGCTGCTGATTTATCTTCGATACAAATGCCCA 1417
 QY 1474 GATATGATCGTCTTTTCAAGTGGCCACTGTTCAATCCAGCTTGTGTTTCCCTTCACATGC 1533
 DB 1418 GATATGATCGTCTTTTCAAGTGGCCACTGTTCAATCCAGCTTGTGTTTCCCTTCACATGC 1477
 QY 1534 CTCTTCATGTTGGCTTCCCTTATTCGGACCCATTTAGTACAGGGATTTGGCTTCGTC 1593
 DB 1478 CTCTTCATGTTGGCTTCCCTTATTCGGACCCATTTAGTACAGGGATTTGGCTTCGTC 1537
 QY 1594 ATCACTCTGACTGGAGTCCCTGCGTATTTATCTCTTATATGCGAAGAAACCCAGG 1653
 DB 1538 ATCACTCTGACTGGAGTCCCTGCGTATTTATCTCTTATATGCGAAGAAACCCAGG 1597
 QY 1654 TGGTTTGAATAATGTCAGAGAAATAACCAAGACATTTACAAATAATCTGGAAGTTGTA 1713
 DB 1598 TGGTTTGAATAATGTCAGAGAAATAACCAAGACATTTACAAATAATCTGGAAGTTGTA 1657
 QY 1714 CCAGAAGAGATAAGTTATGAACCTTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGA 1773
 DB 1658 CCAGAAGAGATAAGTTATGAACCTTAATGGACTTGAGATCTTGGCAATCTGCCCAAGGGA 1717
 QY 1774 GACACAAATAGGGATTTTACTTCTGAAAGTCTAGAGAATTTACAACCTTTGGTG 1833
 DB 1718 GACACAAATAGGGATTTTACTTCTGAAAGTCTAGAGAATTTACAACCTTTGGTG 1777
 QY 1834 ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTTAGCANTATTCGAACCTAATTTCT 1893
 DB 1778 ATAAACAAAAGGAGTCAGTTATTTTATTCATATATTTTAGCANTATTCGAACCTAATTTCT 1837
 QY 1894 AAGAAATTTAGTTATTAACCTATGATGTTATAGAAAGTGAATATGCACTTATTTCTATGAG 1953
 DB 1838 AAGAAATTTAGTTATTAACCTATGATGTTATAGAAAGTGAATATGCACTTATTTCTATGAG 1897
 QY 1954 TCGCACAATTTCTGAGTCTCTGATACCTACCTATTTGGGGTTAGGAGAAAGACTAGACAA 2013
 DB 1898 TCGCACAATTTCTGAGTCTCTGATACCTACCTATTTGGGGTTAGGAGAAAGACTAGACAA 1957
 QY 2014 TTACTATGGTCAATTTCTTACACATATGTTAGCAGCGCAAGAACTTCAAAATTTGAAG 2073
 DB 1958 TTACTATGGTCAATTTCTTACACATATGTTAGCAGCGCAAGAACTTCAAAATTTGAAG 2017
 QY 2074 ACTGAGATTTTCTGTATATATGGGTTTGTAAAGATGGTTTACACACTACAGATGCT 2133
 DB 2018 ACTGAGATTTTCTGTATATATGGGTTTGTAAAGATGGTTTACACACTACAGATGCT 2077
 QY 2134 ATACTGTGAAAGTGTGTTTCAATTCGAAAGAAAGCAATCATGATTTATGCGAAGAG 2193
 DB 2078 ATACTGTGAAAGTGTGTTTCAATTCGAAAGAAAGCAATCATGATTTATGCGAAGAG 2137
 QY 2194 GAGAGAA 2200
 DB 2138 GAGAGAA 2144

RESULT 3
 ABA09201
 ID ABA09201 standard; cdna; 1958 BP.

XX AC ABA09201;
 XX DT 11-JAN-2002 (first entry)
 XX DE Human cysteine/Glu transporter homologue cdna, SEQ ID NO:977.
 XX KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
 KW cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 XX antifungal; vulnery; antiulcer; ss.
 OS Homo sapiens.
 XX WO200157188-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US03800.
 XX 03-FEB-2000; 2000US-0496914.
 XX 27-APR-2000; 2000US-0560875.
 XX (HYSE-) HYSEQ INC.
 XX Tang Yt, Liu C, Drmanac RT;
 XX WPI: 2001-457740/49.
 XX P-PSDB; ABB11957.
 XX Human proteins and DNA encoding sequences useful for preventing,
 XX treating or ameliorating a medical condition in a mammalian subject
 XX e.g. arthritis and cancer -
 XX Claim 1; Page 835; 1963pp; English.
 XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 XX invention also relates to vectors and recombinant host cells comprising a
 XX nucleotide of the invention, methods of producing the novel polypeptides,
 XX antibodies against the polypeptides, methods of detecting the nucleotides
 XX or polypeptides in a sample, and methods of identifying compounds which
 XX bind to polypeptides of the invention. Although novel, many of the
 XX polypeptides of the invention have homology to known proteins, and hence
 XX giving an insight into their probable biological activities, and hence
 XX potential therapeutic applications. The polypeptides of the invention may
 XX have various activities, including cytokine, stem cell growth factor activity;
 XX differentiation activities; receptor or ligand activities; or may be
 XX haematopoiesis regulatory activity; tissue growth activity;
 XX immunomodulatory activity; activin- or inhibin-related activities;
 XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
 XX thrombolytic activities; cancer cell proliferation or metastasis.
 XX involved in oncogenesis, receptor or ligand activities and nucleotides of
 XX the invention are useful for preventing, treating or ameliorating medical
 XX conditions, e.g., by protein or gene therapy. Such conditions include
 XX cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
 XX disorders), chronic inflammatory conditions (e.g. asthma or arthritis),
 XX proliferative retinopathy, atherosclerosis, coronary heart disease,
 XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 XX vascular growth. Polypeptides involved with tissue regeneration and
 XX repair (or nucleic acids encoding them) may be used to promote wound
 XX healing (e.g., of burns, incisions and ulcers), while those with
 XX immunomodulatory activities may be used in the treatment of viral,

Db 1801 TTCTCTGAAAGCTGAGAGAAATACAACCTTTGGTGATATAAACAAGAGCTCACTTATTTTA 1860
 QY 1861 TTCTATATTTTGTAGCATATTCGAACATAATTTCTTAAGAAATTTAGTTAFAACTCTATGTAG 1920
 Db 1861 TTCTATATTTTGTAGCATATTCGAACATAATTTCTTAAGAAATTTAGTTAFAACTCTATGTAG 1920
 QY 1921 TTATAGAAAGTGAATATGAGTATTTCTATGAGTGCACAAATTTCTGAGTCTCTGTATACC 1980
 Db 1921 TTATAGAAAGTGAATATGAGTATTTCTATGAGTGCACAAATTTCTGAGTCTCTGTATACC 1980
 QY 1981 TACCTATTGGGTTAGGAGAAAGACTAGACAATTTACTATGTGTCTATCTTCTACAACAT 2040
 Db 1981 TACCTATTGGGTTAGGAGAAAGACTAGACAATTTACTATGTGTCTATCTTCTACAACAT 2040
 QY 2041 ATGTTAGCACGCAAGAACCTTCAAAATGGAAGACTGAGATTTTTCGTATATATGGGTT 2100
 Db 2041 ATGTTAGCACGCAAGAACCTTCAAAATGGAAGACTGAGATTTTTCGTATATATGGGTT 2100
 QY 2101 TTGTAAAGATGGTTTACACACTACAGATGCTATACTGTGAAAAGTGTTCAAATTCG 2160
 Db 2101 TTGTAAAGATGGTTTACACACTACAGATGCTATACTGTGAAAAGTGTTCAAATTCG 2160
 QY 2161 AAAAAAGCATACATCATGATTATGCAAGAGAGAGAGAGAGTGTGTTTCAATTCG 2220
 Db 2161 AAAAAAGCATACATCATGATTATGCAAGAGAGAGAGAGAGTGTGTTTCAATTCG 2220
 QY 2221 TATTAATAAAAAA 2239
 Db 2221 TATTAATAAAAAA 2239

RESULT 2

AAD23461
 ID AAD23461 standard; cdna; 5981 BP.
 XX
 AC AAD23461;
 XX
 DT 26-FEB-2002 (first entry)
 XX
 DE Human lung tumour-specific 14F10 full length cdna.
 XX
 KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
 KW antisense-therapy; vaccine; immune response; lung cancer; 14F10; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200172295-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 28-MAR-2001; 2001WO-US09991.
 XX
 PR 29-MAR-2000; 2000US-0538037.
 PR 05-JUN-2000; 2000US-0588937.
 PR 18-AUG-2000; 2000US-0640878.
 PR 22-SEP-2000; 2000US-234517P.
 PR 01-NOV-2000; 2000US-0704512.
 PR 14-DEC-2000; 2000US-0738973.
 XX

(CORI-) CORIXA CORP.

XX
 PI Reed SG, Lodes MJ, Mohamath R, Secret H, Benson DR, Indirias CY;
 PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
 XX
 XX WPI; 2001-639201/73.

XX New human lung-specific polynucleotides and polypeptides for the
 PT diagnosis and treatment of disease e.g. lung cancer -
 XX
 XX Claim 1; Page 332-334; 378pp; English.

XX The invention relates to isolated lung tumour-specific proteins and
 CC their corresponding cdna molecules. Lung tumour-specific proteins and

CC their antigen-presenting cells are useful for stimulating and/or
 CC expanding T cells specific for a tumour protein, and for inhibiting
 CC the development of cancer. The invention also relates to a composition
 CC useful for stimulating an immune response, and for treating cancer. The
 CC lung tumour specific oligonucleotide is useful in gene therapy and for
 CC diagnosis, detection and treatment of lung cancer. The present sequence
 CC is a cdna encoding human lung tumour-specific protein.
 XX

SQ Sequence 5981 BP; 1858 A; 1029 C; 1098 G; 1996 T; 0 other;

Query Match 83.4%; Score 1867; DB 23; Length 5981;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 334 GAGAAAGTGCAGCTGAAGAGAGAAAGTCACTTTTACTGAGGGAGTCTCCATTATCATTTGGC 393
 Db 278 GAGAAAGTGCAGCTGAAGAGAGAAAGTCACTTTTACTGAGGGAGTCTCCATTATCATTTGGC 337
 QY 394 ACCATCATTTGGAGCAGGAATCTTCACTCTCTCTAAGGGCGTCTCCAGAACACGGGCAGC 453
 Db 338 ACCATCATTTGGAGCAGGAATCTTCACTCTCTCTAAGGGCGTCTCCAGAACACGGGCAGC 397
 QY 454 GTGGGCATGCTCTGACCATCTGACGGTGTGGGGTCTCTGTCACATTTTGGAGCTTTG 513
 Db 398 GTGGGCATGCTCTGACCATCTGACGGTGTGGGGTCTCTGTCACATTTTGGAGCTTTG 457
 QY 514 TCTTATGCTGAATTTGGGACAACTATATAAAGAAATCTGGAGGTCATTACACATATATTTTG 573
 Db 458 TCTTATGCTGAATTTGGGACAACTATATAAAGAAATCTGGAGGTCATTACACATATATTTTG 517
 QY 574 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGTGGAACTCCTCATATACGC 633
 Db 518 GAAGTCTTTGGTCCATTACAGCTTTTGTACGAGTCTGGTGGAACTCCTCATATACGC 577
 QY 634 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGACGCTCATCTCGAACCATTTTT 693
 Db 578 CCTGCAGCTACTGCTGTGATATCCCTGGCATTTTGACGCTCATCTCGAACCATTTTT 637
 QY 694 ATTCATGTGAATCCTCGAATTCGATCAAGCTCATTTACAGCTGTGGGCATCAACTGTA 753
 Db 638 ATTCATGTGAATCCTCGAATTCGATCAAGCTCATTTACAGCTGTGGGCATCAACTGTA 697
 QY 754 GTGATGGTCTCTAAATAGCATGATGTCAGCTGGAGCGCCCGGATCCAGATTTCTTAAC 813
 Db 698 GTGATGGTCTCTAAATAGCATGATGTCAGCTGGAGCGCCCGGATCCAGATTTCTTAAC 757
 QY 814 TTTTGAAGCTCACAGCAATTTGATAATTTATAGTCCCTGGAGTTATGCAGCTAATTA 873
 Db 758 TTTTGAAGCTCACAGCAATTTGATAATTTATAGTCCCTGGAGTTATGCAGCTAATTA 817
 QY 874 GGTCAACGCGAGAACTTTAAGACGCCCTTTTCAGAGAGAGATTCAAGTATTACGCGTTG 933
 Db 818 GGTCAACGCGAGAACTTTAAGACGCCCTTTTCAGAGAGAGATTCAAGTATTACGCGTTG 877
 QY 934 CCACCTGGCTTTTATTTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGT 993
 Db 878 CCACCTGGCTTTTATTTATGGAATGTATGCATATGCTGGCTGGTTTACCTCAACTTTGT 937
 QY 994 ACTGAAGAAGTAGAAAACCCCTGAAAACCAATTCCTTGCATATGATATATCCATGGCC 1053
 Db 938 ACTGAAGAAGTAGAAAACCCCTGAAAACCAATTCCTTGCATATGATATATCCATGGCC 997
 QY 1054 ATTTGTCACATTTGGCTATGTGCTGACAAATGTGCCCTACTTTAGGACCATTAATGCTGAG 1113
 Db 998 ATTTGTCACATTTGGCTATGTGCTGACAAATGTGCCCTACTTTAGGACCATTAATGCTGAG 1057
 QY 1114 GAGCTGCTGCTTTCAAAATGAGTGGCAGTCACTTTTCTGAGCGGCTACTGGGAAATTC 1173
 Db 1058 GAGCTGCTGCTTTCAAAATGAGTGGCAGTCACTTTTCTGAGCGGCTACTGGGAAATTC 1117
 QY 1174 TCATTAGCAGTTCGATCTTTGTTGCTCTCTCTGCTTTGGCTTCCATGAACGGTGGTGTG 1233
 Db 1118 TCATTAGCAGTTCGATCTTTGTTGCTCTCTCTGCTTTGGCTTCCATGAACGGTGGTGTG 1177

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 09:16:28 ; Search time 522 Seconds
(without alignments)
9659.433 Million cell updates/sec

Title: US-09-667-170A-440
Perfect score: 2239
Sequence: 1 ggaagtgaagtgcagag.....ttattataaaaaaaaaa 2239

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N:Genesec_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2239	100.0	2239	23	Human lung tumour-
2	1867	83.4	5981	23	Human lung tumour-
3	1719	76.8	1958	22	Human cysteine/Glu
4	1719	76.8	1958	22	Human EST-derived
5	414	18.5	772	20	Human gene express
6	300	13.4	300	20	Human gene express
7	300	13.4	300	20	Human gene express
8	288	12.9	288	20	Human gene express
9	284	12.7	300	21	Human colon cancer

C	10	281	12.6	337	23	AAD23462	Human lung tumour-
	11	269	12.0	596	24	AAK27651	Human colon cancer
	12	108	4.8	800	20	AAZ16609	Human gene express
	13	107	4.8	575	24	ABL37718	Human colon tumour
	14	100	4.5	123	22	AAK68639	Human immune/haema
	15	60	2.7	60	24	ABN35211	Human spliced tran
	16	51	2.3	761	20	AAZ17094	Human gene express
	17	45	2.0	983	20	AAZ53014	Human prostate tum
	18	33	1.5	136	22	AAK82863	Human immune/haema
	19	33	1.5	136	22	AAK82863	Human immune/haema
	20	33	1.5	8948	22	AAK67209	Human immune/haema
	21	33	1.5	175737	24	ABK83571	Human CDNA differe
C	22	32	1.4	300	20	AAZ13802	Human gene express
	23	32	1.4	300	20	AAZ13190	Human gene express
	24	32	1.4	2440	24	ABL64408	Stomach cancer rel
C	25	32	1.4	3403	22	AAH18197	Human CDNA sequenc
C	26	32	1.4	8894	23	ABK42719	Genomic sequence #
	27	31	1.4	300	20	AAZ14169	Human gene express
	28	31	1.4	1443	22	AAZ17386	Human bone marrow
	29	30	1.3	356	24	ABL83406	Human ovarian canc
C	30	30	1.3	569	22	AAH06560	Human CDNA clone (
C	31	30	1.3	804	22	AAH16361	Human CDNA sequenc
C	32	30	1.3	2246	22	AAH16361	Human nervous syst
	33	30	1.3	25012	22	ABA15431	Human immune/haema
	34	30	1.3	31804	22	AAK90698	Human immune/haema
	35	30	1.3	38894	22	AAK69848	Human ovarian canc
C	36	29	1.3	482	24	ABL87814	Human ovarian canc
C	37	29	1.3	492	24	ABL82980	Human immune/haema
C	38	29	1.3	501	24	ABL80020	Human immune/haema
C	39	28	1.3	17674	24	ABL33345	Human immune syste
	40	27	1.2	359	23	ABV01115	Human prostate exp
	41	27	1.2	360	23	ABV10284	Human prostate exp
	42	27	1.2	407	23	ABV31455	Human prostate exp
	43	27	1.2	407	23	ABV40424	Human transcriptio
	44	27	1.2	737	22	AAK02169	Human nervous syst
	45	27	1.2	3637	22	ABA20309	

ALIGNMENTS

RESULT 1
AAD23460
ID AAD23460 standard; cDNA; 2239 BP.
XX
AC AAD23460;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human lung tumour-specific 19A4 cDNA.
XX
KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; 19A4; ss.
XX
OS Homo sapiens.
XX
PN WO200172295-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US09991.
XX
PR 29-MAR-2000; 2000US-0538037.
PR 05-JUN-2000; 2000US-0588937.
PR 18-AUG-2000; 2000US-0640878.
PR 22-SEP-2000; 2000US-2345178.
PR 01-NOV-2000; 2000US-0704512.
PR 14-DEC-2000; 2000US-0738973.
XX
(CORI-) CORIXA CORP.
PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;

Thu Apr 17 09:48:35 2003

PRIOR APPLICATION NUMBER: US 60/357,600
 PRIOR FILING DATE: 2002-02-15
 NUMBER OF SEQ ID NOS: 54
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 9
 LENGTH: 6296
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-163-866-9

Query Match 16.9%; Score 379.2; DB 9; Length 6296;

Best Local Similarity 55.9%; Pred. No. 4.1e-91;
 Matches 720; Conservative 0; Mismatches 568; Indels 0; Gaps 0;

QY	334	GAGAAAGTGCAGCTGAAGAGGAAGTCACTTTACTCAGGGAGTCTCCATTATCATTTGGC	393
DB	370	GAAACTATGACAGTGAAGAGGAGATCTCCCTGCTGAATGGGGTCAAGCTGGTGGGC	429
QY	394	ACCATCATTTGGAGCAGGAATCTTCATCTCTCCTTAAGGGCTGCTCCAGAACACGGGACG	453
DB	430	AACATGATCGGCTCAGGATCTTTGCTCACCAAGGGTGTGCTGTACACACTGCCCTC	489
QY	454	GTGGGCATGCTCTGACCATCTGGACGGTGTGGGGTCTCTGTCATATTGGAGCTTTG	513
DB	490	TATGGGATGTCACTGATTTGTGGGCAATTTGGTGGCTCTTCTCTGTTGGTGGCTT	549
QY	514	TCTTATGCTGAATTTGGGAACAACCTATAAAGAAATCTGGAGTCAATTTACACATATATTG	573
DB	550	TGTTATGACAGAGTGGGACCAACCATCAAGTCCGGAGCCAGCTACGCTTATATTTCTA	609
QY	574	GAAGTCTTTGGTGCATTTACAGCTTTTGTAGGAGTCTGGGTGGAACCTCCTCATATAATACG	633
DB	610	GAGGCTTTGGGGCTTCATTTGCCCTCATCGGCTGTGGGTCTCAGCTGCTAGTTGTTGAG	669
QY	634	CTGTGAGTCTGCTGTGATATCTCTGCAATTTGGAGCTTACATTTCTGGAACCAATTTT	693
DB	670	CCCACGGTCAAGGCATCATCGCCATCACCTTTGCCAATACATCATCCAGCGCTCTTC	729
QY	694	ATTCAATGTGAATCCCTGAACTTTGGATCAAGCTCATTTACAGCTGTGGCATAACTGTA	753
DB	730	CCAGCTGTGATCCCCATACCTGGCCCTGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTG	789
QY	754	GTGATGCTCTAAATAGCATGATGCTCAGCTGAGCGCCCGATCCAGATTTCTTTAAAC	813
DB	790	CTGACATTTGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	849
QY	814	TTTTCGAAGCTCACAGCAATTTCTGATAATATATAGTCCCTGGAGTTATGACAGTAAATAA	873
DB	850	TACGCCAAGGTCGTAGCGCTCATTTGCCATCATTTGTATGGGCTTTGTAAACTGTGCCAG	909
QY	874	GGTCAACGCAGAACCTTTAAAGACGCTTTTTCAGGAAGAGATTTCAAGTATTTACGCGTTG	933
DB	910	GGACACTCTGAGCACTTTTCAGGACGCTTTGAGGGTTCCTCTGGACATGGGAAACCTC	969
QY	934	CCACTGGCTTTTATTTATGGAATGTATGATATGCTGGCTGGTGTCTTACTCAACTTTGTT	993
DB	970	TCCTTGCCCTCTACTCTGCTCTCTCTTACTCAGGTGGGACACCCCTTAATTTGTA	1029
QY	994	ACTGAAGAAGTAGAAACCTGAAACCAATCCCTTGCATATATGATATATATATATATATAT	1053
DB	1030	ACAGAGAAATCAAAACCCAGAAAGAAATTTGCCCTTGGCCATTTGGGATTTCTATGCCA	1089
QY	1054	ATTCTGACCATTTGGCTATGCTGACAAATGTGGCCCTACTTTACGACCAATTAATGCTGAG	1113
DB	1090	ATTGTGAGCTCATCTACATCTCTGACCAATGTGGCCCTATTACACAGTGTGAACATTTCA	1149
QY	1114	GAGCTGCTGCTTTTCAATAGCAGTGGGACCTTTTCTGAGGGCTACTGGGAAATTC	1173
DB	1150	GATGCTCTTAGCAGTATGCTGTGGTGTGACATTTGCTGACACAGAGCTTTGGCATGTTTC	1209
QY	1174	TCATTAGCAGTTCGATCTTTGTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1233
DB	1210	AGCTGGACATCCCATTTGCTGTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1269

Search completed: April 16, 2003, 10:55:37
 Job time : 241 secs

QY	1234	TTTGTCTCTCCAGGTTATTTCTATGTTGCGTCTCGAGAGGTCACCTTCAGAAATCCTC	1293
DB	1270	TTTGTCTCATCAAGGTTGTTCTGTTGGCTCCCGGGAGGCGCACCTACCGGACCTTCTG	1329
QY	1294	TCCATGATTCATGTCGCAAGCACACTCTCTACAGCTGTTATTTGTTGACCCCTTTG	1353
DB	1330	TCCATGATCCACATTTGAGCGTTTACACCTATCCCTGCTTTACTGTTCAATTCACCATG	1389
QY	1354	ACAATGATAATGCTCTTCTCTGGAGACCTCGACAGTCTTTTGAATTTCTCAGTTTGGCC	1413
DB	1390	GCACTCATCTACCTCATCTGTTGAGGATGTTTCCAGCTTATCAACTACTTTCAGCTTCAGC	1449
QY	1414	AGTGGCTTTTATTTGGGCTGGCAGTTGCTGGCTGATTTATCTTCGATACAAATGCCCCA	1473
DB	1450	TACTGTGTTCTTCTGTTGGGCTGCTGTTGTTGGACAGCTCTACCTCCGCTGGAAGAGGCC	1509
QY	1474	GATATGATCTGCTCTTTCAAGGTGCCACTGTTCATCCAGCTTTTGTTCCTTCACATGC	1533
DB	1510	AAGCGGCCCGGCTCTCAAGCTGAGCGTGTTTTCCCATCGTGTCTGCTATGCTCTC	1569
QY	1534	CTCTTCATGTTGCCCTTTCCCTCTATTCCGACCCATTTTAGTACAGGAGTTGGCTTCGTC	1593
DB	1570	GTGTTCTGTTGATAGTGGCCCTTTCACTGACACCAATTAATTCCTCATTTGGCATCGGG	1629
QY	1594	ATCACTCTGACTGGAGTCCCTGCGTATT	1621
DB	1630	ATTGCCCTTTTGGAGTCCCTTTCTACT	1657

Query Match	16.9%;	Score 379.2;	DB 9;	Length 6295;
Best Local Similarity	55.9%;	Pred. No. 4.1e-91;		
Matches '720;	Conservative	0;	Mismatches 568;	Indels 0;
	0;			
334	GAGAAAGTCGACGTGAAGAGGAAAGTCTACTTTACTGAGGGAGTCTCCATATCATTTGCG	393		
DB				
363	GAAACTATGACGTGAAGAGAGAGATCTCCCTGCTGAATGGGTGACCTGGTGGTGCGC	422		
DB				
394	ACCATATTGGACGGAATCTTCATCTCTCTAAGGGCGTGTCCAGAACACGGCGACG	453		
DB				
423	AACATGATCGGTCACGGGATCTTGTCTCACCCAAAGGGTGTGTGTACACACTGCCTCC	482		
DB				
454	GTGGCAGTCTCTGACCATCTGGACGTTGTGGGGTCTCTGTCATATTGGAGCTTTG	513		
DB				
483	TATGGGATCTCACTGATTGTGGGCCATTGGTGGGCTCTTCTCTTTGGGTGCCCTT	542		
DB				
514	TCATTATGCTGAATTGGGACAACTATTAAGAAATCTGGAGGTCAATTACACATATATTTG	573		
DB				
543	TGTTATGACAGCTGGGACCAACCATCACAAGTCGGGAGCCAGCTACGTTATATCTTA	602		
DB				
574	GAAGTCTTTGGTCCATTACCAGCTTTTGTACGAGTCTGGGTGGAAGTCTCTCAATAAGC	633		
DB				
603	GAGGCGCTTTGGGGGCTTCATTGCCCTTCATCCGCTGTGGGTCTCACTGCTAGTTGTTGAG	662		
DB				
634	CCTGACGTACTGCTGTGATATCCCTGGCATTTTGGACGCTACATTTGGAACCAATTTT	693		
DB				
663	CCCACGGGTACGGGCATCATCGCCATCACCTTTTGCCAACTACATCATCCAGCGCTCCTC	722		
DB				
694	ATTCAATGTGAATCCCTGAACTTGGATCAGCTCAATTACAGCTGTGGGCATAACTGTA	753		
DB				
723	CCGAGCTGTGATCCCCATACCTGGGCTGCCGCTCTCCTGGCTGCTTGCATATGTCGTG	782		
DB				
754	GTGATGGTCTTAATAGCATGAGTGTGTCAGCTGGAGGCCCGGATCCAGATTTCTTAAAC	813		
DB				
783	CTGACATTTTGGAAGTGTGGCTATGCAAGTGGGGCACACGTGTGCAGGACACGCTTCACT	842		
DB				
814	TTTTGCAAGCTCACAGCAATTCGATAATATAGTCCCTGGAGTTATGACAGCTAATTTAA	873		
DB				

[illegible]

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-53

Query Match 46.8%; Score 1047.4; DB 9; Length 1268;
Best Local Similarity 87.9%; Pred. No. 1.7e-271;
Matches 1238; Conservative 0; Mismatches 1; Indels 170; Gaps 2;

QY 182 GAAAGCCTGTTGTCTCCACATCTCCAAAGGAGGTACCTGCAGGGAATGTTAACGGGA 241
Db |||||||
QY 242 GCGTGCCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCCCTTTTCAGGAAGAGACG 301
Db |||||||
QY 302 CCTTTTCAGGAAGAGACGCCCTTTTCAGGAAGAGAGAGTGCAGCTGAAGAGGAAAGTCA 361
Db |||||||
QY 362 CTTTACTGAGGGAGTCTCCATTTATCATTTGGCACCATTATGGAGCAAGGAATCTTCATCT 421
Db |||||||
QY 422 CTCTTAAGGGCGTCTCCAGAACACGGCAGCGTGGGCATGCTCTGCACCATCTTCATCT 215
Db |||||||
QY 216 CTCTTAAGGGCGTCTCCAGAACACGGCAGCGTGGGCATGCTCTGCACCATCTTCATCT 275
QY 482 TGTGGGGTCTCTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACCTATAA 541
Db |||||||
QY 276 TGTGGGGTCTCTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACCTATAA 335
QY 542 AGAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTGGTCCATTACCAGCTTTTG 601
Db |||||||
QY 336 AGAATCTGGAGGTCATTACACATATATTTTGAAGTCTTTGGTCCATTACCAGCTTTTG 395
QY 602 TACAGTCTGGTGGAACTCCTCATATACGCCCTCAGCTACTGCTGTGATATCCCTGG 455
Db |||||||
QY 396 TACAGTCTGGTGGAACTCCTCATATACGCCCTCAGCTACTGCTGTGATATCCCTGG 455
QY 662 CATTTGGACCTCATCTCTGGAACCAATTTTATCAATGTGAAATCCCTGAACTTGCGA 721
Db |||||||
QY 456 CATTTGGACCTCATCTCTGGAACCAATTTTATCAATGTGAAATCCCTGAACTTGCGA 515
QY 722 TCAAGTCTATTACAGTCTGGGCAATCACTGATGATGCTCTTAATAGCATGAGTGCA 781
Db |||||||
QY 516 TCAAGTCTATTACAGTCTGGGCAATCACTGATGATGCTCTTAATAGCATGAGTGCA 575
QY 782 GCTGGAGCGCCGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 841
Db |||||||
QY 576 GCTGGAGCGCCGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATAA 635
QY 842 TTATAGTCCCTGGAGTTATGCACTAATTAAGGTCACAGCAATTTTGAAGCTCACAGCA 901
Db |||||||
QY 636 TTATAGTCCCTGGAGTTATGCACTAATTAAGGTCACAGCAATTTTGAAGCTCACAGCA 695
QY 902 TTTTCAAGAGATTCAGATTTACGCGTTGCCACTGGCTTTTATATGGAATGATG 961
Db |||||||
QY 696 TTTTCAAGAGATTCAGATTTACGCGTTGCCACTGGCTTTTATATGGAATGATG 755
QY 962 CATATGCTGGCTGTTTACCTCACTTTGTTACTGAAGAGTAGAAGCAATTTGCTGACAA 1021
Db |||||||
QY 756 CATATGCTGGCTGTTTACCTCACTTTGTTACTGAAGAGTAGAAGCAATTTGCTGACAA 815
QY 1022 CCATTCCCTTGCATATGATATCCATGGCCATTGTGACCATTTGCTGACAA 1081
Db |||||||
QY 816 CCATTCCCTTGCATATGATATCCATGGCCATTGTGACCATTTGCTGACAA 875
QY 1082 ATGTGGCTCTTACGACCAATTAATGCTGAGGAGTGTGCTTTCAATGCAAGTGGCAG 1141
Db |||||||
QY 876 ATGTGGCTCTTACGACCAATTAATGCTGAGGAGTGTGCTTTCAATGCAAGTGGCAG 935
QY 1142 TGACCTTTTCTAGCGGCTACTGGAAATTTCTCATTAGCAGTTCGATCTTTGTGGCC 1201

Db 936 TGACCTTTTCTGAGCGGTACHTGGAAATTTCTCATTAGCAGTTCCGATCTTTGTTGCC- 994
QY 1202 TCTCTGCTTTGGCTCCATGAACGGTGTGTGTTGTCTGCTCCAGGTTATTTCTATGTTG 1261
Db ----- 994
QY 1262 CGTCTCGAGAGGTTCACCTTCCAGAAATCCCTCCATGATTCATGTCGGAAGCACACTC 1321
Db -----CCCTC 999
QY 1322 CTCTACCAAGCTTTATTTGTTTTCACCCCTTTGACAAATGATAATGCTCTCTCTGGAGACC 1381
Db |||||||
QY 1000 CTCTACCAAGCTTTATTTGTTTGCACCTTTGACAAATGATAATGCTCTCTCTGGAGACC 1059
QY 1382 TCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTTGGGCTGGCAGTTG 1441
Db |||||||
QY 1060 TCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTTGGGCTGGCAGTTG 1119
QY 1442 CTGGGCTGATTTATCTTTCGATACAAATGCCAGATATGCATCTCTTTCAAGGTGCCAC 1501
Db |||||||
QY 1120 CTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCTCTTTCAAGGTGCCAC 1179
QY 1502 TGTTCATCCCAAGCTTTCTTCTTCCATGCTCTCTTCAAGGTGCCCTTTCCCTCTATT 1561
Db |||||||
QY 1180 TGTTCATCCCAAGCTTTCTTCTTCCATGCTCTCTTCAAGGTGCCCTTTCCCTCTATT 1239
QY 1562 CGGACCAATTTAGTACAGGATTTGGCTTC 1590
Db |||||||
QY 1240 CGGACCAATTTAGTACAGGATTTGGCTTC 1268

RESULT 12

US-10-163-866-32
; Sequence 32, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLIC7S AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; PRIOR APPLICATION NUMBER: 2002-06-05
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-32

Query Match

Best Local Similarity 20.1%; Score 449; DB 9; Length 520;
Matches 513; Conservative 0; Mismatches 0; Indels 54; Gaps 1;
QY 182 GAAAGCCTGTTGTCTCCACATCTCCAAAGGAGGTACCTGCAGGGAATGTTAACGGGA 241
Db |||||||
QY 8 GAAAGCCTGTTGTCTCCACATCTCCAAAGGAGGTACCTGCAGGGAATGTTAACGGGA 67
QY 242 GCGTGCCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCCCTTTTCAGGAAGAGACG 301
Db |||||||
QY 68 GCGTGCCTTCCCTGGGCAACAGGAGCCACCTGGGAG----- 105
QY 302 CCTTTTCAGGAAGAGAGCGCTTTTTCAGGAAGAGAGAAAGTGCAGTGAAGAGAAAGTCA 361

Thu Apr 17 09:48:35 2003

;; PRIORITY FILING DATE: 2001-10-22
;; PRIORITY APPLICATION NUMBER: US 60/357,253
;; PRIORITY FILING DATE: 2002-02-15
;; PRIORITY APPLICATION NUMBER: US 60/357,600
;; PRIORITY FILING DATE: 2002-02-15
;; NUMBER OF SEQ ID NOS: 54
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 34
;; LENGTH: 2000
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-163-866-34

Query Match 61.3%; Score 1373.4; DB 9; Length 2000;

Best Local Similarity 96.0%; Pred. No. 0;

Matches 1441; Conservative 0; Mismatches 6; Indels 54; Gaps 1;

QY 182 GAAAGGCTGTGTGTCCACCATCTCCAAAGGAGGTTACTGCGAGGGAATGTTAAACGGGA 241
DB 143 GAAAGGCTGTGTGTCCACCATCTCCAAAGGAGGTTACTGCGAGGGAATGTTAAACGGGA 202
QY 242 GGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCCCTTTTCAGGAAGAGAGACG 301
DB 203 GGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGAGGAGCCCTTTTCAGGAAGAGAGACG 240
QY 302 CTTTTCAGGAAG 240
DB 241 -----GAGAAAGTGCAGCTGAAGAGGAAGTCA 361
QY 362 CTTTACTGAGGGAGTCTCATTTATCATTTGGCACCACATATTGGAGCAGGAATCTTTCATCT 421
DB 269 CTTTACTGAGGGAGTCTCATTTATCATTTGGCACCACATATTGGAGCAGGAATCTTTCATCT 328
QY 422 CTCCTAAGGGGCTGCTCCAGAACACGGGAGCGTGGGCAATGCTCTGACCATCTTGAGCGG 481
DB 329 CTCCTAAGGGGCTGCTCCAGAACACGGGAGCGTGGGCAATGCTCTGACCATCTTGAGCGG 388
QY 482 TGTGTGGGGTCTGTCTACTATTGAGAGCTTTGCTTATGCTGAATTTGGGAACAACATATAA 541
DB 389 TGTGTGGGGTCTGTCTACTATTGAGAGCTTTGCTTATGCTGAATTTGGGAACAACATATAA 448
QY 542 AGAAATCTGAGGTCATTACACATATATTTTGGAAAGCTTTTGGTCCATACAGCTTTTG 601
DB 449 AGAAATCTGAGGTCATTACACATATATTTTGGAAAGCTTTTGGTCCATACAGCTTTTG 508
QY 602 TACGAGTCTGGTGAATCTCTCATATACCCCTGAGCTACTGCTGTGATATCCCTGG 661
DB 509 TACGAGTCTGGTGAATCTCTCATATACCCCTGAGCTACTGCTGTGATATCCCTGG 568
QY 662 CATTTGGAGCGCTACATTTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTTGGGA 721
DB 569 CATTTGGAGCGCTACATTTCTGGAACCATTTTATTCAATGTGAATCCCTGAACCTTGGGA 628
QY 722 TCAAGCTCATTTACAGCTGAGGATACCTAGTGTGATGAGTCTTAATAGATGATGATCA 781
DB 629 TCAAGCTCATTTACAGCTGAGGATACCTAGTGTGATGAGTCTTAATAGATGATGATCA 688
QY 782 GCTGAGAGCGCGGATCCAGATTTCTTAAACCTTTTGAAGCTCACAGCAATCTGATATA 841
DB 689 GCTGAGAGCGCGGATCCAGATTTCTTAAACCTTTTGAAGCTCACAGCAATCTGATATA 748
QY 842 TTATAGTCCCTGGAGTTATGAGCTAAATTAAGGTCAACGAGCAATTTAAAGAGCGCT 901
DB 749 TTATAGTCCCTGGAGTTATGAGCTAAATTAAGGTCAACGAGCAATTTAAAGAGCGCT 808
QY 902 TTTTCAAGGAGAGATTTCAAGTATTACCGGTTGCCACTGCTTTTATTATGGAATGTATG 961
DB 809 TTTTCAAGGAGAGATTTCAAGTATTACCGGTTGCCACTGCTTTTATTATGGAATGTATG 868
QY 962 CATATGCTGGCTGTTTACCTCAACTTTGTTACTGAAGAGTACAGAAACCTGAAAAA 1021
DB 869 CATATGCTGGCTGTTTACCTCAACTTTGTTACTGAAGAGTACAGAAACCTGAAAAA 928

QY 1022 CCATTCCCTTGGCAATATATATATATCCATGCGCCATTTGTCACCATTTGCTATGTCGTGACAA 1081
DB 929 CCATTCCCTTGGCAATATATATATATATCCATGCGCCATTTGTCACCATTTGCTATGTCGTGACAA 988
QY 1082 ATGTGGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTTTCAAAATGCAAGTGGCAG 1141
DB 989 ATGTGGGCTACTTTACGACCAATTAATGCTGAGGAGCTGCTTTCAAAATGCAAGTGGCAG 1048
QY 1142 TCACCTTTTCTGAGCGGCTACTGGAATTTCTCATTTAGCAGTTCCGATTTTGTGTCGCC 1201
DB 1049 TGACCTTTTCTGAGCGGCTACTGGAATTTCTCATTTAGCAGTTCCGATTTTGTGTCGCC 1108
QY 1202 TCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGCTTATCTATGTTG 1261
DB 1109 TCTCTGCTTTGGCTCCATGAACGGTGGTGTGTTGCTGCTCCAGCTTATCTATGTTG 1168
QY 1262 CGTCTGAGAGGCTCACCTTCCAGAAATCCCTCCATGATTCATGTCGCGCAAGCACACTC 1321
DB 1169 CGTCTGAGAGGCTCACCTTCCAGAAATCCCTCCATGATTCATGTCGCGCAAGCACACTC 1228
QY 1322 CTCCTACAGCTGTTATTTGTTTTGTCACCCCTTTTGACAATGATAATGCTCTTCTTGAGACC 1381
DB 1229 CTCCTACAGCTGTTATTTGTTTTGTCACCCCTTTTGACAATGATAATGCTCTTCTTGAGACC 1288
QY 1382 TCGACAGCTTTTGAATTTCCCTCAGTTTGGCAGCTGGCTTTTATTTGGCTGGCAGTTG 1441
DB 1289 TCGACAGCTTTTGAATTTCCCTCAGTTTGGCAGCTGGCTTTTATTTGGCTGGCAGTTG 1348
QY 1442 CTGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTTCAAGGTGGCAC 1501
DB 1349 CTGGCTGATTTATCTTCGATACAAATGCCAGATATGATCGTCTTCAAGGTGGCAC 1408
QY 1502 TGTTCATPCCAGCTTTGTTTTTCCCTTCCATGCTCTTTCATGCTTCCCTCTTAT 1561
DB 1409 TGTTCATPCCAGCTTTGTTTTTCCCTTCCATGCTCTTTCATGCTTCCCTCTTAT 1468
QY 1562 GCGACCCATTTAGTACAGGATTTGGCTTCTCATCTACTCTGAGTGGAGTCCCTGGGTATT 1621
DB 1469 GCGACCCATTTAGTACAGGATTTGGCTTCTCATCTACTCTGAGTGGAGTCCCTGGGTATT 1528
QY 1622 ATCTCTTTATATATGAGGAGAAAGAACCCAGGTGGTTTGAATAATGTCAGAGAAATAA 1681
DB 1529 ATCTCTTTATATATGAGGAGAAAGAACCCAGGTGGTTTGAATAATGTCAGAGTTCCTAG 1588
QY 1682 C 1682
DB 1589 C 1589

RESULT 11

US-10-163-866-53
; Sequence 53, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 1268

; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-163-866-52

Query Match 63.9%; Score 1431.8; DB 9; Length 1528;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1497; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY	182	GAAGCCCTGTTGTGTCACCATCTCCAAAGGAGGTTACCTGCAGGAAATGTTAAAGGGA	241
DB	30	GAAGCCCTGTTGTGTCACCATCTCCAAAGGAGGTTACCTGCAGGAAATGTTAAAGGGA	89
QY	242	GGCTGCCTTCCCTGGGCAAGAGCCACCTGGGAGGAGCGCTTTTCAGGAAGAGAGC	301
DB	90	GGCTGCCTTCCCTGGGCAAGAGGAGCCACCTGGGAGC-----	127
QY	302	CTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAGTGCAGCTGAAGAGGAAGTCA	361
DB	128	-----GAGAAAGTGCAGCTGAAGAGGAAGTCA	155
QY	362	CTTTACTGAGGGAGTCTCATATCATTTGGACACCATCAATTGGACAGGAATCTTCATCT	421
DB	156	CTTTACTGAGGGAGTCTCATATCATTTGGACACCATCAATTGGACAGGAATCTTCATCT	215
QY	422	CTCCTAAGGGCGTCTCCAGAACACAGGGCAGCGTGGGATGTCCTGACCATCTGGAGG	481
DB	216	CTCCTAAGGGCGTCTCCAGAACACAGGGCAGCGTGGGATGTCCTGACCATCTGGAGG	275
QY	482	TGTTGGGTCCTGTCACTATTGGAGCTTTGCTTTATGCTGAATTTGGAAACAATAA	541
DB	276	TGTTGGGTCCTGTCACTATTGGAGCTTTGCTTTATGCTGAATTTGGAAACAATAA	335
QY	542	AGAAATCTGGAGGTCATTACACATATTTTGGAAAGTCTTTGGTCCATTACCGCTTTG	601
DB	336	AGAAATCTGGAGGTCATTACACATATTTTGGAAAGTCTTTGGTCCATTACCGCTTTG	395
QY	602	TACGAGTCTGGGTGAACCTCTCATATACCGCTTGCAGCTACTGCTGTGATATCCCTGG	661
DB	396	TACGAGTCTGGGTGAACCTCTCATATACCGCTTGCAGCTACTGCTGTGATATCCCTGG	455
QY	662	CATTTGGAGCTACATTTCTGGAACCATTTTATTCAATGGAATCCCTGAACTTGGGA	721
DB	456	CATTTGGAGCTACATTTCTGGAACCATTTTATTCAATGGAATCCCTGAACTTGGGA	515
QY	722	TCAAGCTCATACAGCTGTGGGCACTAGTGTAGTGTGCTTAAATAGCATGAGTGTCA	781
DB	516	TCAAGCTCATACAGCTGTGGGCACTAGTGTAGTGTGCTTAAATAGCATGAGTGTCA	575
QY	782	GCTGAGGCGCCGGATCCAGATTTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATAA	841
DB	576	GCTGAGGCGCCGGATCCAGATTTCTTAACCTTTTGGCAAGCTCACAGCAATTTCTGATAA	635
QY	842	TTATAGTCCCTGGAGTTATGAGCTAATAAAGGTCAACCGAGAACTTTAAAGAGCGCT	901
DB	636	TTATAGTCCCTGGAGTTATGAGCTAATAAAGGTCAACCGAGAACTTTAAAGAGCGCT	695

QY	902	TTTCAGGAAGAGATTCAAGTATTACCGGTTGCCACTGGCTTTTATATGAATGTATG	961
DB	696	TTTCAGGAAGAGATTCAAGTATTACCGGTTGCCACTGGCTTTTATATGAATGTATG	755
QY	962	CATATGCTGGCTGGTTTACCTCAACTTTTACTGTAAGAGTGAAGAAACCCCTGAAAAA	1021
DB	756	CATATGCTGGCTGGTTTACCTCAACTTTTACTGTAAGAGTGAAGAAACCCCTGAAAAA	815
QY	1022	CCATTCCCTTGCATATGATATCCATGCGCAATTTGCAATTTGCTGCTGCTGCTGCTG	1081
DB	816	CCATTCCCTTGCATATGATATCCATGCGCAATTTGCAATTTGCTGCTGCTGCTGCTG	875
QY	1082	ATGTGGCTTACTTACGACCAATTAATGCTGAGGAGTGTGCTTCAAAATGCAAGTGGCAG	1141
DB	876	ATGTGGCTTACTTACGACCAATTAATGCTGAGGAGTGTGCTTCAAAATGCAAGTGGCAG	935
QY	1142	TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTTCCGATCTTTGTTGCC	1201
DB	936	TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTTCCGATCTTTGTTGCC	995
QY	1202	TCTCTGCTTTGGCTCCATGAAGGGTGTGTTGCTGCTCCAGTTTATCTATGTTG	1261
DB	996	TCTCTGCTTTGGCTCCATGAAGGGTGTGTTGCTGCTCCAGTTTATCTATGTTG	1055
QY	1262	GGTCTGAGAGGGTACCTTCCAGAAATCTCTCCATGATTCATGTCGCAAGACACTC	1321
DB	1056	GGTCTGAGAGGGTACCTTCCAGAAATCTCTCCATGATTCATGTCGCAAGACACTC	1115
QY	1322	CTCTACCACTGTTATGTTTTCACCCCTTTGACAATGATAATGCTCTCTGGAGACC	1381
DB	1116	CTCTACCACTGTTATGTTTTCACCCCTTTGACAATGATAATGCTCTCTGGAGACC	1175
QY	1382	TCGACAGTCTTTTGAATTTCCCTCAGTTTTCAGAGTGGCTTTTATGGGCTGGCAGTTG	1441
DB	1176	TCGACAGTCTTTTGAATTTCCCTCAGTTTTCAGAGTGGCTTTTATGGGCTGGCAGTTG	1235
QY	1442	CTGGGCTGATTTATCTTCGATACAAATGCCAGATATCATCGTCTTCAAGTGGCAC	1501
DB	1236	CTGGGCTGATTTATCTTCGATACAAATGCCAGATATCATCGTCTTCAAGTGGCAC	1295
QY	1502	TGTTATCCACAGCTTTGTTTTCCTTCACATGCTTTCATGCTGCTGCTGCTGCTGCTGCT	1561
DB	1296	TGTTATCCACAGCTTTGTTTTCCTTCACATGCTTTCATGCTGCTGCTGCTGCTGCTGCT	1355
QY	1562	CGGACCCATTAGTACAGGATTTGCTTCATCACTCTGCTGCTGCTGCTGCTGCTGCTGCT	1621
DB	1356	CGGACCCATTAGTACAGGATTTGCTTCATCACTCTGCTGCTGCTGCTGCTGCTGCTGCT	1415
QY	1622	ATCTCTTTTATATATGGGACAAAGAACCCAGTGGTTAGTAATGTCAGAGAAATAA	1681
DB	1416	ATCTCTTTTATATATGGGACAAAGAACCCAGTGGTTAGTAATGTCAGAGAAATAA	1475
QY	1682	CCAGAACATTACAATAATATCTGGAAGTTGTACCAAGAGAGATAAGTTATGA	1734
DB	1476	CCAGAACATTACAATAATATCTGGAAGTTGTACCAAGAGAGATAAGTTATGA	1528

RESULT 10

US-10-163-866-34
; Sequence 34, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733

Thu Apr 17 09:48:35 2003

```

US-10-163-866-33
; Sequence 33, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 33
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-163-866-33

Query Match      65.0%; Score 1454.4; DB 9; Length 1542;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1519; Conservative 1; Mismatches 54; Gaps 1;

QY 182 GAAAGCTGTGTGTCACCATCTCCAAAGGAGGTTTACCTGCAGGGAAATGTTAACGGGA 241
DB 16 GAAAGCTGTGTGTCACCATCTCCAAAGGAGGTTTACCTGCAGGGAAATGTTAACGGGA 75
QY 242 GGCTGCTCCCTCGGGCAACAGGACCTGGGCGAGGACGCTTTTCAGGAAGAGAG 301
DB 76 GGCTGCTCCCTCGGGCAACAGGAGGACCTGGGCGAG----- 113
QY 302 CCTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
DB 114 -----GAGAAAGTGCAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 141
QY 362 CTTTACTGAGGGAGTCTCCATATCATTTGGACCATCATTTGAGCAGGAAATCTTCATCT 421
DB 142 CTTTACTGAGGGAGTCTCCATATCATTTGGACCATCATTTGAGCAGGAAATCTTCATCT 201
QY 422 CTCCTAAGGGCGTCTCCAGAACACGGCGAGCGTGGGCGATGCTCTGACCATCTGGACGG 481
DB 202 CTCCTAAGGGCGTCTCCAGAACACGGCGAGCGTGGGCGATGCTCTGACCATCTGGACGG 261
QY 482 TGTGTGGGGTCTGTCATATTTGGAGCTTTGCTTATGCTGAATGGGAACAACATATA 541
DB 262 TGTGTGGGGTCTGTCATATTTGGAGCTTTGCTTATGCTGAATGGGAACAACATATA 321
QY 542 AGAAATCTGGAGTCAATACACATATATTTGGAGTCTTTGGTCCATTAACAGCTTTTG 601
DB 322 AGAAATCTGGAGTCAATACACATATATTTGGAGTCTTTGGTCCATTAACAGCTTTTG 381
QY 602 TAGAGTCTGGGTGGAACTCCATATAATACGCCCTGGAGCTACTGCTGTGATATCCCTGG 661
DB 382 TAGAGTCTGGGTGGAACTCCATATAATACGCCCTGGAGCTACTGCTGTGATATCCCTGG 441
QY 662 CATTGTGAGCTACATCTGGAAACATTTTATTCATATGTAATCCCTGAATCTGGGA 721
DB 442 CATTGTGAGCTACATCTGGAAACATTTTATTCATATGTAATCCCTGAATCTGGGA 501
QY 722 TCAAGCTCAATACAGCTGTGGGCAATACATGATGCTGATGCTGATGCTGATGCTGATGCT 781
DB 502 TCAAGCTCAATACAGCTGTGGGCAATACATGATGCTGATGCTGATGCTGATGCTGATGCT 561
QY 782 GCTGGAGCGCGGATCCAGATTTCTTACCTTTTGGCAAGCTTCACAGCAATTTCTGATA 841

```

```

DB 562 GCTGGAGCGCCCGGATCCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATA 621
QY 842 TTATAGTCCCTGGAGTTATGCAGCTAAATAAGGTCACAGCAACTTTAAAGACGCT 901
DB 522 TTATAGTCCCTGGAGTTATGCAGCTAAATAAGGTCACAGCAACTTTAAAGACGCT 681
QY 902 TTTTCAGGAAGAGATTTCAAGTATTTACGCGTTCGCACTGGCTTTTATTTATGGAATGATG 961
DB 682 TTTTCAGGAAGAGATTTCAAGTATTTACGCGTTCGCACTGGCTTTTATTTATGGAATGATG 741
QY 962 CATATGCTGGCTGTTTACCTCAACTTTTACTGAGCAAGTACAGCAACTTTTAAAGACGCT 1021
DB 742 CATATGCTGGCTGTTTACCTCAACTTTTACTGAGCAAGTACAGCAACTTTTAAAGACGCT 801
QY 1022 CCATTTCCCTTGCATATATATTCATGAGCAATTTCTCATTTAGCAGTTCCGATTTTGTGGCC 1081
DB 802 CCATTTCCCTTGCATATATATTCATGAGCAATTTCTCATTTAGCAGTTCCGATTTTGTGGCC 861
QY 1082 ATGTGGCTGCTTACGACCAATTAATCTGAGGAGCTGGCTTTTCAAAATGCAAGTGGCAG 1141
DB 862 ATGTGGCTGCTTACGACCAATTAATCTGAGGAGCTGGCTTTTCAAAATGCAAGTGGCAG 921
QY 1142 TGACCTTTTTCAGCGGCTACTGGCAATTTCTCATTTAGCAGTTCCGATTTTGTGGCC 1201
DB 922 TGACCTTTTTCAGCGGCTACTGGCAATTTCTCATTTAGCAGTTCCGATTTTGTGGCC 981
QY 1202 TCTCTGCTTTGGCTTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTTATTTCTATGTTG 1261
DB 982 TCTCTGCTTTGGCTTCCATGAACGGTGGTGTGTTGCTGCTCCAGGTTATTTCTATGTTG 1041
QY 1262 CGTCTGAGAGGCTCAGCTTCCAGAAATCTCTCATGATTCATGTCGCAAGCACTC 1321
DB 1042 CGTCTGAGAGGCTCAGCTTCCAGAAATCTCTCATGATTCATGTCGCAAGCACTC 1101
QY 1322 CTCTACAGCTGTATTTGTTTGGACCCCTTTCACCAATGATGATGCTCTCTCGAGACC 1381
DB 1102 CTCTACAGCTGTATTTGTTTGGACCCCTTTCACCAATGATGATGCTCTCTCGAGACC 1161
QY 1382 TCAGACAGCTTTTGAATTTCCCTCAGTTTTCAGAGTGGCTTTTATTTGGGCTGGCAGTTG 1441
DB 1162 TCAGACAGCTTTTGAATTTCCCTCAGTTTTCAGAGTGGCTTTTATTTGGGCTGGCAGTTG 1221
QY 1442 CTGGGCTGATTTATCTTCGATACAAATGCGCAGATGATGCTCTCTTCAAGGTGGCCAC 1501
DB 1222 CTGGGCTGATTTATCTTCGATACAAATGCGCAGATGATGCTCTTCAAGGTGGCCAC 1281
QY 1502 TGTTCATCCAGCTTTGTTTTCCTTTCACATGCTCTTCATGTTGCTTCCCTTCCCTTAT 1561
DB 1282 TGTTCATCCAGCTTTGTTTTCCTTTCACATGCTCTTCATGTTGCTTCCCTTCCCTTAT 1341
QY 1562 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCTGCTGAGTGGCTCCCTGCTGAT 1621
DB 1342 CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCTGCTGAGTGGCTCCCTGCTGAT 1401
QY 1622 ATCTCTTTTATATGAGCAAGAAACCCAGGTGTTTGAATAGTAAATGTCAGAGAAATAA 1681
DB 1402 ATCTCTTTTATATGAGCAAGAAACCCAGGTGTTTGAATAGTAAATGTCAGAGAAATAA 1461
QY 1682 CCAGAACATTAACAATAACTGGAAGTTGTACCAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
DB 1462 CCAGAACATTAACAATAACTGGAAGTTGTACCAGAGAGAGAGAGAGAGAGAGAGAGAG 1521
QY 1742 GACTTGGAGATCTGG 1755
DB 1522 GACTTGGAGATCTGG 1535

```

RESULT 9
 US-10-163-866-52
 ; Sequence 52, Application US/10163866
 ; Publication No. US20030027188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.

RESULT 7

US-10-163-866-30
 ; Sequence 30, Application US/10163866
 ; Publication No. US20030027186A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.
 ; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
 ; FILE REFERENCE: EX02-080C
 ; CURRENT APPLICATION NUMBER: US/10/163,866
 ; PRIOR FILING DATE: 2002-06-05
 ; PRIOR FILING DATE: 2002-06-05
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR FILING DATE: 2002-02-15
 ; PRIOR FILING DATE: 2002-02-15
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 30
 ; LENGTH: 1861
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-163-866-30

Query Match 68.8%; Score 1540.8; DB 9; Length 1861;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1606; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

QY	182	GAAGCCCTGTTGTCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAAGCGGA	241
DB	243	GAAGCCCTGTTGTCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAAGCGGA	302
QY	242	GGCTGCCTTCCCTGGGCAACAAAGGAGCCACCTCGGCGAGGACGCCCTTTTCAGGAAGAGAG	301
DB	303	GGCTGCCTTCCCTGGGCAACAAAGGAGCCACCTGGGCGAG	340
QY	302	CTTTTCAGGAAGAGAGCGCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGAGGAAGTCA	361
DB	341	-----GAGAAAGTGCAGCTGAAGAGGAAGTCA	368
QY	362	CTTTACTAGGGGAGTCTCATTTATCATTTGGCACCACATCTTGGCAGAGAAATCTTCATCT	421
DB	369	CTTTACTAGGGGAGTCTCATTTATCATTTGGCACCACATCTTGGCAGAGAAATCTTCATCT	428
QY	422	CTCCTAAGGCGTCTCCAGACACAGCGGCGTGGGCGATCTCTGACCATCTGGAGG	481
DB	429	CTCCTAAGGCGTCTCCAGACACAGCGGCGTGGGCGATCTCTGACCATCTGGAGG	488
QY	482	TGRTGGGTCTGTCTACATTTATTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAA	541
DB	489	TGTGTGGGTCTGTCTACATTTATTTGGAGCTTTTGTCTTATGCTGAATTTGGGAACAACATAA	548
QY	542	AGAAATCTGGAGTCTTACACATATATTTTGAAGCTTTTGGTCCATTTACAGCTTTTG	601
DB	549	AGAAATCTGGAGTCTTACACATATATTTTGAAGCTTTTGGTCCATTTACAGCTTTTG	608
QY	602	TACAGTCTGGGTGGAATCTCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG	661
DB	609	TACAGTCTGGGTGGAATCTCTCATATAATACGCCCTGCAGCTACTGCTGTGATATCCCTGG	668
QY	662	CATTTGGAGCGTACATCTGGAACCATTTTATTCATGTAATGAAATCCCTGAACTTGGCA	721
DB	669	CATTTGGAGCGTACATCTGGAACCATTTTATTCATGTAATGAAATCCCTGAACTTGGCA	728
QY	722	TCAAGCTCATTTACAGCTGTGGGATTAACCTAGTGTGCTTAAATAGCATGAGTGCA	781
DB	729	TCAAGCTCATTTACAGCTGTGGGATTAACCTAGTGTGCTTAAATAGCATGAGTGCA	788
QY	782	GCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTTCACAGCAATTTCTGATAA	841

DB	789	GCTGGAGCGCCGGATCCAGATTTTCTTAACCTTTTGAAGCTTCACAGCAATTTCTGATAA	848
QY	842	TTATAGTCCCTGGAGTATGACAGTAATTAAGGTCAAAACGAGAGACTTTAAAGACGCT	901
DB	849	TTATAGTCCCTGGAGTATGACAGTAATTAAGGTCAAAACGAGAGACTTTAAAGACGCT	908
QY	902	TTTCAGAGAGATTCAGATTAATGACGGTTCGCCACTGGCTTTTATTAATGGAATGATG	961
DB	909	TTTCAGAGAGATTCAGATTAATGACGGTTCGCCACTGGCTTTTATTAATGGAATGATG	968
QY	962	CATATGCTGCTGGTGTACCTCAACTTGTACTGAAGAAGTAGAAAAACCTGAAAAA	1021
DB	969	CATATGCTGCTGGTGTACCTCAACTTGTACTGAAGAAGTAGAAAAACCTGAAAAA	1028
QY	1022	CCATTTCCCTTGCATATGATATATCATGAGGAGTGTGCTTCAAAATGCAAGTGGCAG	1081
DB	1029	CCATTTCCCTTGCATATGATATATCATGAGGAGTGTGCTTCAAAATGCAAGTGGCAG	1088
QY	1082	ATGTGCGCTACTTTACGACCATTAATGCTGAGGAGTGTGCTTCAAAATGCAAGTGGCAG	1141
DB	1089	ATGTGCGCTACTTTACGACCATTAATGCTGAGGAGTGTGCTTCAAAATGCAAGTGGCAG	1148
QY	1142	TGACCTTTTCTGAGCGCTACTGGAAATTTCTCATAGCAGTTCGATCTTTGTTGCC	1201
DB	1149	TGACCTTTTCTGAGCGCTACTGGAAATTTCTCATAGCAGTTCGATCTTTGTTGCC	1208
QY	1202	TCTCTGCTTGGCTCCATGAAAGGAGTGTGCTTCAAAATGCAAGTGGCAGTTCGATCTTTGTTGCC	1261
DB	1209	TCTCTGCTTGGCTCCATGAAAGGAGTGTGCTTCAAAATGCAAGTGGCAGTTCGATCTTTGTTGCC	1268
QY	1262	CGTCTCAGAGGAGTACCTTCCAGAAATCTCTCCATGATTCATGTCGCCAGACACATC	1321
DB	1269	CGTCTCAGAGGAGTACCTTCCAGAAATCTCTCCATGATTCATGTCGCCAGACACATC	1328
QY	1322	CTCTACAGCTGTTATTTGTTTGCACCTTTGACAAATGATGCTCTCTCGAGAC	1381
DB	1329	CTCTACAGCTGTTATTTGTTTGCACCTTTGACAAATGATGCTCTCTCGAGAC	1388
QY	1382	TCGACAGCTTTTGAATTTCCAGTTTCCAGAGTGGCTTTTATTTGGGTGGCAGTTG	1441
DB	1389	TCGACAGCTTTTGAATTTCCAGTTTCCAGAGTGGCTTTTATTTGGGTGGCAGTTG	1448
QY	1442	CTGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGCCAC	1501
DB	1449	CTGGCTGATTTATCTTCGATACAAATGCCAGATATGCATCGTCTTCAAGGTGCCAC	1508
QY	1502	TGTTCACTCCAGCTTTGTTTCTTCACATGCTCTTCATGCTTCCCTTTCCCTCTATT	1561
DB	1509	TGTTCACTCCAGCTTTGTTTCTTCACATGCTCTTCATGCTTCCCTTTCCCTCTATT	1568
QY	1562	CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACGTGGAGTCCCTGCTATT	1621
DB	1569	CGGACCCATTTAGTACAGGATTTGGCTTCGTCATCACTCTGACGTGGAGTCCCTGCTATT	1628
QY	1622	ATCTCTTTATTTATGAGCAAGAAACCCAGGTGGTTTACAATATGTCAGAGAAATAA	1681
DB	1629	ATCTCTTTATTTATGAGCAAGAAACCCAGGTGGTTTACAATATGTCAGAGAAATAA	1688
QY	1682	CCAGAACATTTACAATTAATCTGGAAGTTGTACAGAGAAAGATAAGTTATGAACATAAG	1741
DB	1689	CCAGAACATTTACAATTAATCTGGAAGTTGTACAGAGAAAGATAAGTTATGAACATAAG	1748
QY	1742	GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGATTTTACTTCAAT	1801
DB	1749	GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATAGGATTTTACTTCAAT	1808
QY	1802	TTCTGAAAGCTAGAGATTTACAACTTTGGTGATAAACAATAA	1843
DB	1809	TTCTGAAAGCTAGAGATTTACAACTTTGGTGATAAACAATAA	1850

RESULT 8

```

; LENGTH: 2482
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-29

Query Match      84.3%; Score 1888.2; DB 9; Length 2482;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1965; Conservative      3; Indels 55; Gaps

QY 182 GAAAGCGCTGTGTGTCACCATCTCCAAAGGAGGTTTACCTGCAGGGAATGTTAACGGGA 241
DB 239 GAAAGCGCTGTGTGTCACCATCTCCAAAGGAGGTTTACCTGCAGGGAATGTTAACGGGA 298
QY 242 GCGTGCCTTCCTCTGGGCAACAGGAGCCACCTCGGGCAGGACGCCCTTTTCAGGAAGAGAG 301
DB 299 GCGTGCCTTCCTCTGGGCAACAGGAGCCACCTCGGGCAG ----- 336
QY 302 CCTTTTCAGGAAGAGAGCGCCCTTTTCAGGAAGAGAGAGAGTGCAGCTGAAGAGGAAGTCA 361
DB 337 -----GAGAAAGTGCAGCTGAAGAGGAAGTCA 364
QY 362 CTTTACTCAGGGAGTCTCCATTATCATTTGGCACCATCATTTGAGCAGGAAATCTTCATCT 421
DB 365 CTTTACTCAGGGAGTCTCCATTATCATTTGGCACCATCATTTGAGCAGGAAATCTTCATCT 424
QY 422 CTCCTTAAGGGCGTGTCCAGAACAGGCGACGTGGGCATCTCTGCACCATCTGCACGG 481
DB 425 CTCCTTAAGGGCGTGTCCAGAACAGGCGACGTGGGCATCTCTGCACCATCTGCACGG 484
QY 482 TCGTGTGGGTCCTGTCTACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAA 541
DB 485 TATGTGGGGTCTGTCTACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAACAACATAA 544
QY 542 AGAAATCTGGAGGTCAATTACACATATATTTTGGAGTCTTTTGGTCCATTAACAGCTTTTG 601
DB 545 AGAAATCTGGAGGTCAATTACACATATATTTTGGAGTCTTTTGGTCCATTAACAGCTTTTG 604
QY 602 TAGAGTCTGGGTGGAACCTCCATTAATACGCCCTGACGCTACTGCTGTGTATATCCCTGG 661
DB 605 TAGAGTCTGGGTGGAACCTCCATTAATACGCCCTGACGCTACTGCTGTGTATATCCCTGG 664
QY 662 CATTTGGAGCGCTACATCTCGGAACCATTTTTTATTCAATGTGAATCCCTGAACTTCGGA 721
DB 665 CATTTGGAGCGCTACATCTCGGAACCATTTTTTATTCAATGTGAATCCCTGAACTTCGGA 724
QY 722 TCAAGCTCATTTACAGCTGTGGGCAATACCTGTAGTGATGGTCTTAATAGCATGAGTGCA 781
DB 725 TCAAGCTCATTTACAGCTGTGGGCAATACCTGTAGTGATGGTCTTAATAGCATGAGTGCA 784
QY 782 GCTGGAGCGCCGGATCCAGATTTCTTAAACCTTTTTCGAAGCTCACAGCAATCTTGATAA 841
DB 785 GCTGGAGCGCCGGATCCAGATTTCTTAAACCTTTTTCGAAGCTCACAGCAATCTTGATAA 844
QY 842 TTATAGTCCCTGGAGTTATCGACCTAATTAAGGTCAAACGAGAACTTTTAAAGACGCT 901
DB 845 TTATAGTCCCTGGAGTTATCGACCTAATTAAGGTCAAACGAGAACTTTTAAAGACGCT 904
QY 902 TTTTCAGGAAGAGATTCAAGTATTACGGGTTGGCAGCTGGCTTTTATTATGGAATGTATG 961
DB 905 TTTTCAGGAAGAGATTCAAGTATTACGGGTTGGCAGCTGGCTTTTATTATGGAATGTATG 964
QY 962 CATATGCTGGCTGGTTTTTACCTCAACTTTGTACTGAGAAGTAGAAAAACCCCTGAAAAAA 1021
DB 965 CATATGCTGGCTGGTTTTTACCTCAACTTTGTACTGANGAGTAGAAAAACCCCTGAAAAA 1024
QY 1022 CCATTTCCCTTGCATATGTATATCCCATGGCCATTGTCCACCATTTGGCTATGTCTGCACAA 1081
DB 1025 CCATTTCCCTTGCATATGTATATCCCATGGCCATTGTCCACCATTTGGCTATGTCTGCACAA 1084
QY 1082 ATGTGGCTACTTTTACGACCATTAATGCTGAGAGAGCTGCTCTTTCAAAATGCAGTGGCAG 1141
DB 1085 ATGTGGCTACTTTTACGACCATTAATGCTGAGAGAGCTGCTCTTTCAAAATGCAGTGGCAG 1144

```

Db 134 CTTTACTGAGGGAGTCTCCATTATCATTTGGCACCATCATTTGGAGCAGGAATCTTCACT 193
Qy 422 CTCCTAAGGGCTGCTCCAGAACACGGGAGCGTGGCGATGCTCTGACCATCTGGACGG 481
Db 194 CTCCTAAGGGCTGCTCCAGAACACGGGAGCGTGGCGATGCTCTGACCATCTGGACGG 253
Qy 482 TGTGTGGGGTCCCTGCTCACTATTTGGAGCTTTGCTCTATCTGAAATGGGAACAACATAAA 541
Db 254 TGTGTGGGGTCCCTGCTCACTATTTGGAGCTTTGCTCTATCTGAAATGGGAACAACATAAA 313
Qy 542 AGAAATCTGAGGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATACACAGCTTTTG 601
Db 314 AGAAATCTGAGGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATACACAGCTTTTG 373
Qy 602 TAGAGTCTGGGTGGAACTCCCTCAATAATACGCCCTGACGCTACTGCTGTGATATCCCTGG 661
Db 374 TAGAGTCTGGGTGGAACTCCCTCAATAATACGCCCTGACGCTACTGCTGTGATATCCCTGG 433
Qy 662 CATTTGGAGCTACATCTCGGAACCATTTTATTCATATGTGAATCCCTGAACTTCGGA 721
Db 434 CATTTGGAGCTACATCTCGGAACCATTTTATTCATATGTGAATCCCTGAACTTCGGA 493
Qy 722 TCAAGTCTAATACGCTGAGGCACTAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 781
Db 494 TCAAGTCTAATACGCTGAGGCACTAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 553
Qy 782 GCTGAGGCGCGGATCCAGATTTTCTTAACCTTTTGAAGCTCAGACGATTCGATGA 841
Db 554 GCTGAGGCGCGGATCCAGATTTTCTTAACCTTTTGAAGCTCAGACGATTCGATGA 613
Qy 842 TTATAGTCCCTGGAGTTATGAGCTAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 901
Db 614 TTATAGTCCCTGGAGTTATGAGCTAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 733
Qy 902 TTTGAGGAGAGATTCAGATTTACGCGGTTTGCACCTTGGCTTTTATGATGATGATG 961
Db 674 TTTGAGGAGAGATTCAGATTTACGCGGTTTGCACCTTGGCTTTTATGATGATGATG 733
Qy 962 CATATGCTGGCTGCTTTTACCTCAACTTTGTTACTGAAGAGTGAAGAACCTCGAAAAA 1021
Db 734 CATATGCTGGCTGCTTTTACCTCAACTTTGTTACTGAAGAGTGAAGAACCTCGAAAAA 793
Qy 1022 CCATTCCCTTGGCAATATGATATCCATGGCCATTTGCTGCTCCAGGTTATTCATGTTG 1081
Db 794 CCATTCCCTTGGCAATATGATATCCATGGCCATTTGCTGCTCCAGGTTATTCATGTTG 853
Qy 1082 ATGTGGCTACTTTAGACCATTAATGCTGAGGAGTCTGCTTTCAATGAGTGGCAG 1141
Db 854 ATGTGGCTACTTTAGACCATTAATGCTGAGGAGTCTGCTTTCAATGAGTGGCAG 913
Qy 1142 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTCCGATCTTTGTTGCC 1201
Db 914 TGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTTAGCAGTCCGATCTTTGTTGCC 973
Qy 1202 TCTCCCTGCTTTGGCTCAATGAGCGTGGTGTGTTGCTGCTCCAGGTTATTCATGTTG 1261
Db 974 TCTCCCTGCTTTGGCTCAATGAGCGTGGTGTGTTGCTGCTCCAGGTTATTCATGTTG 1033
Qy 1262 CGTCTGAGAGGCTCACCTTCAGAAATCCCTCCATGATTCATGTCGCGAAGCAGCACT 1321
Db 1034 CGTCTGAGAGGCTCACCTTCAGAAATCCCTCCATGATTCATGTCGCGAAGCAGCACT 1093
Qy 1322 CTTCTAGAGCTTTATGTTTGGACCCCTTTGACAATGATAATGCTCTCTCTGGAGACC 1381
Db 1094 CTTCTAGAGCTTTATGTTTGGACCCCTTTGACAATGATAATGCTCTCTCTGGAGACC 1153
Qy 1382 TCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTTGGCTGGCAGTTG 1441
Db 1154 TCGACAGCTTTTGAATTTCTCAGTTTGGCAGGTGGCTTTTATTTGGCTGGCAGTTG 1213
Qy 1442 CTGGCTGATTTATCTTCGATACAAATGCCAGATATGATGCTCTCTCAAGGTGCCAC 1501
Db 1214 CTGGCTGATTTATCTTCGATACAAATGCCAGATATGATGCTCTCTCAAGGTGCCAC 1273

RESULT 6

US-10-163-866-29

; Sequence 29, Application US/10163866

; Publication No. US20030027188A1

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: SLIC7s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX02-080C

; CURRENT APPLICATION NUMBER: US/10/163,866

; CURRENT FILING DATE: 2002-06-05

; PRIOR APPLICATION NUMBER: US 60/296,076

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/328,605

; PRIOR FILING DATE: 2001-10-10

; PRIOR APPLICATION NUMBER: US 60/338,733

; PRIOR FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/357,253

; PRIOR FILING DATE: 2002-02-15

; PRIOR APPLICATION NUMBER: US 60/357,600

; PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 29

Qy 1502 TGTTCATCCAGCTTTGTTTTCCTTCACATGCTCTTCATGCTGCCCTTCCCTCTATT 1561
Db 1274 TGTTCATCCAGCTTTGTTTTCCTTCACATGCTCTTCATGCTGCCCTTCCCTCTATT 1333
Qy 1562 CGACCCATTTAGTACAGGATTTGGCTCGTCACTCTGACTGGAGTCCCTGCGTATT 1621
Db 1334 CGACCCATTTAGTACAGGATTTGGCTCGTCACTCTGACTGGAGTCCCTGCGTATT 1393
Qy 1622 ATCTCTTTATATATGGGACAAAGAACCCAGGCTTTAGAAATATATGTCAGAGAAATAA 1681
Db 1394 ATCTCTTTATATATGGGACAAAGAACCCAGGCTTTAGAAATATATGTCAGAGAAATAA 1453
Qy 1682 CCAGAACATTAACAATAATATCTGGAAGTTGTACCAAGAAAGATAAGTTATGAACATAATG 1741
Db 1454 CCAGAACATTAACAATAATATCTGGAAGTTGTACCAAGAAAGATAAGTTATGAACATAATG 1513
Qy 1742 GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATATAGGATTTTACTTCATT 1801
Db 1514 GACTTGAGATCTTGGCAATCTGCCAAGGGGAGACACAAATATAGGATTTTACTTCATT 1573
Qy 1802 TTCTCAAGAGTCTAGAGATTTACAACCTTTGGTGATAAACAAGAGTCACTTATTTTAT 1861
Db 1574 TTCTCAAGAGTCTAGAGATTTACAACCTTTGGTGATAAACAAGAGTCACTTATTTTAT 1633
Qy 1862 TCATATATTTAGCATATTCGAACCTAATTTCTAGAAATTTAGTTATTACTCTATGATG 1921
Db 1634 TCATATATTTAGCATATTCGAACCTAATTTCTAGAAATTTAGTTATTACTCTATGATG 1693
Qy 1922 TATAGAAGTGAATATGAGTTATTCATGAGTCGCACAAATTTCTGAGTCTCTGATACCT 1981
Db 1694 TATAGAAGTGAATATGAGTTATTCATGAGTCGCACAAATTTCTGAGTCTCTGATACCT 1753
Qy 1982 ACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTTGGTCAATCTCTACACATA 2041
Db 1754 ACCTATTTGGGTTAGGAGAAAGACTAGACAAATTTACTATGTTGGTCAATCTCTACACATA 1813
Qy 2042 TGTTCAGCGGCAAGAACTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGSGTTT 2101
Db 1814 TGTTCAGCGGCAAGAACTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGSGTTT 1873
Qy 2102 TGTAAAGATGTTTACACACTACAGATCTCTATACTGTGAAAAGTGTTCATTCCTGA 2161
Db 1874 TGTAAAGATGTTTACACACTACAGATCTCTATACTGTGAAAAGTGTTCATTCCTGA 1933
Qy 2162 AAAAAAGCATACATCATGATTATGCAAGAGAGAGAGAGTA 2204
Db 1934 AAAAAAGCATACATCATGATTATGCAAGAGAGAGAGAGAA 1976

Db 1747 TCTGAAGCTCTAGAGAAATTAACAACCTTTGGTGATAAACAAGAGGAGTCACTATTTATTTATTT 1806
Qy 1863 CATATATTTTACGATATTCGAACCTAAATTTCTTAAGAAATTTAGTTAATAACCTCTATGTAGTT 1922
Db 1807 CATATATTTTACGATATTCGAACCTAAATTTCTTAAGAAATTTAGTTAATAACCTCTATGTAGTT 1866
Qy 1923 ATAGAAAGTGAATATGCAAGTTATTTCTATGAGTGCACAAATTTCTTGAGTCTCTGATACCTA 1982
Db 1867 ATAGAAAGTGAATATGCAAGTTATTTCTATGAGTGCACAAATTTCTTGAGTCTCTGATACCTA 1926
Qy 1983 CCTATTTGGGTTAGAGAAAGACTAGACAATTTACTATGTTGGTCTCATTTCTTACAAACATAT 2042
Db 1927 CCTATTTGGGTTAGAGAAAGACTAGACAATTTACTATGTTGGTCTCATTTCTTACAAACATAT 1986
Qy 2043 GTTACACGCGCAAGAACTTCAAAATTTGAAGACTGAGATTTTCTGTATATATATGGGTTTT 2102
Db 1987 GTTACACGCGCAAGAACTTCAAAATTTGAAGACTGAGATTTTCTGTATATATATGGGTTTT 2046
Qy 2103 GTAAAGATGGTTTACACACTACAGATGTCTATCTACTGTGAAGTGTCTTCAATCTTGAA 2162
Db 2047 GTAAAGATGGTTTACACACTACAGATGTCTATCTACTGTGAAGTGTCTTCAATCTTGAA 2106
Qy 2163 AAAAGCATACATCATGATTATGGCAAGAGGAGAGAGGTA 2204
Db 2107 AAAAGCATACATCATGATTATGGCAAGAGGAGAGAGAA 2148
RESULT 5
US-10-163-866-31
; Sequence 31. Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163,866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338,733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357,600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 31
; LENGTH: 3144
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-163-866-31
Query Match 84.9%; Score 1900.2; DB 9; Length 3144;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1966; Conservative 0; Mismatches 3; Indels 54; Gaps 1;
Qy 182 GAAAGCCTCTGTGTCACCATCTCCAAAGGAGGTTACCTCAGGGAAATTTAACGGGA 241
Db 8 GAAAGCCTCTGTGTCACCATCTCCAAAGGAGGTTACCTCAGGGAAATTTAACGGGA 67
Qy 242 GGCTGCTTCCCTGGGCAACAAGAGCCACCTGGGCGAGGACCCCTTTTCAGGAAGAGACG 301
Db 68 GGCTGCTTCCCTGGGCAACAAGAGCCACCTGGGCGAC----- 105
Qy 302 CCTTTTCAGGAAGAGACGCTTTTCAGGAAGAGAGAGAGTGCAGCTGAAGAGGAAGTCA 361
Db 106 -----GAGAAAGTGCAGCTGAAGAGGAAGTCA 133
Qy 362 CTTTACTGAGGGAGTCTCCATTATTTGGCAACCATCATTTGGCAACCATCATTTGGCAACCATCATTT 421

Db 667 CAAGCTCATATACAGCTGGGCAATAAAGTATGATGCTCTTAATAGCATGAGTGTGAG 726
Qy 783 CTGGAGCGCCGATCCAGATTTCTTAAACCTTTTGAAGCTCACAGCAATTTCTGATAAT 842
Db 727 CTGGAGCGCCGATCCAGATTTCTTAAACCTTTTGAAGCTCACAGCAATTTCTGATAAT 786
Qy 843 TATAGTCCCTGAGATATGCAAGTAAATTAAGGTTCAACGCGAGAACTTTTAAAGACGCTT 902
Db 787 TATAGTCCCTGAGATATGCAAGTAAATTAAGGTTCAACGCGAGAACTTTTAAAGACGCTT 846
Qy 903 TTCAAGAAAGATTTCAAGTATTTACCGGTTGCCACTGGCTTTTATATGAATGTATGC 962
Db 847 TTCAAGAAAGATTTCAAGTATTTACCGGTTGCCACTGGCTTTTATATGAATGTATGC 906
Qy 963 ATATGCTGGCTGGTTTACCTCAACTTTTCTTACTCAAGAGTGAAGAAACCTTGAAAAAC 1022
Db 907 ATATGCTGGCTGGTTTACCTCAACTTTTCTTACTCAAGAGTGAAGAAACCTTGAAAAAC 966
Qy 1023 CATTCCCTTGCATATGATATCCATGGCCATTTGTACCATTTGGCTATGTCTGACAAA 1082
Db 967 CATTCCCTTGCATATGATATCCATGGCCATTTGTACCATTTGGCTATGTCTGACAAA 1026
Qy 1083 TGTGGCTACTTTACAGACCAATTAATGCTGAGGAGTGTCTTCAAAATGCAAGTGGCAGT 1142
Db 1027 TGTGGCTACTTTACAGACCAATTAATGCTGAGGAGTGTCTTCAAAATGCAAGTGGCAGT 1086
Qy 1143 GACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTTCGCACTTTTGTGGCCCT 1202
Db 1087 GACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATAGCAGTTCGCACTTTTGTGGCCCT 1146
Qy 1203 CTCCTGCTTTGGCTCCATGAACGGTGTGTGCTGCTCCAGGTATTTCTATGTTGC 1262
Db 1147 CTCCTGCTTTGGCTCCATGAACGGTGTGTGCTGCTCCAGGTATTTCTATGTTGC 1206
Qy 1263 CTCCTGAGGAGTTCACCTTCCAGAAATTCCTCTCATGATTCATGTCGCAAGCACACTCC 1322
Db 1207 CTCCTGAGGAGTTCACCTTCCAGAAATTCCTCTCATGATTCATGTCGCAAGCACACTCC 1266
Qy 1323 TCTACAGCTGTTATGTTTTCACCTTTGACAAATGATGCTCTCTCTGAGACCT 1382
Db 1267 TCTACAGCTGTTATGTTTTCACCTTTGACAAATGATGCTCTCTCTGAGACCT 1326
Qy 1383 CCACAGCTTTTGAATTTCTCAGTTTTCGAGGTGGCTTTTATTTGGCTGGCAGTTGC 1442
Db 1327 CCACAGCTTTTGAATTTCTCAGTTTTCGAGGTGGCTTTTATTTGGCTGGCAGTTGC 1386
Qy 1443 TGGGCTGATTTATCTTCGATCAAAATGCCAGATATGATGCTCTTCAAGTGGCACT 1502
Db 1387 TGGGCTGATTTATCTTCGATCAAAATGCCAGATATGATGCTCTTCAAGTGGCACT 1446
Qy 1503 GTTCATCCAGCTTTGTTTCTTCCATGCTCTTCAATGTTGGCTTCCCTTCCCTCTATTC 1562
Db 1447 GTTCATCCAGCTTTGTTTCTTCCATGCTCTTCAATGTTGGCTTCCCTTCCCTCTATTC 1506
Qy 1563 GGACCATTTAGTACAGGATTTGGCTTGTCTACTCTGCTGAGTCCCTGGCTATTA 1622
Db 1507 GGACCATTTAGTACAGGATTTGGCTTGTCTACTCTGCTGAGTCCCTGGCTATTA 1566
Qy 1623 TCTCTTTATATATGGCAAGAAACCCAGGTGTTTGAATAATATGTCAGAGAAATAAC 1682
Db 1567 TCTCTTTATATATGGCAAGAAACCCAGGTGTTTGAATAATATGTCAGAGAAATAAC 1626
Qy 1683 CAGAACATTAACAATAATATGGAAGTTGTACAGAAAGATAAGTTATGAACATAATGG 1742
Db 1627 CAGAACATTAACAATAATATGGAAGTTGTACAGAAAGATAAGTTATGAACATAATGG 1686
Qy 1743 ACTTGAGATCTTGCCCAATGCGCCAGGGAGACACAAATAGGATTTTACTTCATTT 1802
Db 1687 ACTTGAGATCTTGCCCAATGCGCCAGGGAGACACAAATAGGATTTTACTTCATTT 1746
Qy 1803 TCTCAAGTCTAGAGAAATTAACAATTTGTGTGATTAACAAGAGGTCAGTTATTTTAT 1862

1801 TTCTGAAAGTCTAGAGAAATACAACTTTGGTGATAAACAAGAGAGTCAAGTTATTTTA 1860
1801 TTCTGAAAGTCTAGAGAAATACAACTTTGGTGATAAACAAGAGAGTCAAGTTATTTTA 1860
1861 TTCTGAAAGTCTAGAGAAATACAACTTTGGTGATAAACAAGAGAGTCAAGTTATTTTA 1920
1861 TTCTGAAAGTCTAGAGAAATACAACTTTGGTGATAAACAAGAGAGTCAAGTTATTTTA 1920
1921 TTATGAAAGTCTAGAGAAATACAACTTTGGTGATAAACAAGAGAGTCAAGTTATTTTA 1980
1921 TTATGAAAGTCTAGAGAAATACAACTTTGGTGATAAACAAGAGAGTCAAGTTATTTTA 1980
1981 TACCTATTGGGGTTAGGAGAAAAGACTAGACAAATTTACTATGTGGTCAATTTCTGATATG 2040
1981 TACCTATTGGGGTTAGGAGAAAAGACTAGACAAATTTACTATGTGGTCAATTTCTGATATG 2040
2041 ATGTTAGCAGCGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGTT 2100
2041 ATGTTAGCAGCGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGGTT 2100
2101 TTGTAAGAGATGTTTACACACTACAGATGCTATACACTGTGAAAAGTGTTCATTCCTG 2160
2101 TTGTAAGAGATGTTTACACACTACAGATGCTATACACTGTGAAAAGTGTTCATTCCTG 2160
2161 AAAAAAGCATACATCATGATTATGCAAGAGGAGAGAGTGTGTTTCTTAAAT 2220
2161 AAAAAAGCATACATCATGATTATGCAAGAGGAGAGAGTGTGTTTCTTAAAT 2220
2221 TATTAATAAAAAAAAAAAAAA 2239
2221 TATTAATAAAAAAAAAAAAAA 2239

RESULT 3
US-09-854-133-441
; Sequence 441, Application US/09854133
; Publication No. US2002018349A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 441
; LENGTH: 5981
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-133-441

Query Match 92.9%; Score 2080.8; DB 9; Length 5981;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 2146; Conservative 0; Mismatches 2; Indels 54; Gaps 1;

3 AGGTTGAAGTGAGCAGATCATGCCAGCTGGGTGACAGTGAGACTCTGTCTCAAAACAG 62
1 AGGTTGAAGTGAGCAGATCATGCCAGCTGGGTGACAGTGAGACTCTGTCTCAAAACAG 60
63 AATTAGGAAAAAGAAAAAGAAAAAGAGAGAGAGAGAAATTCACGCCAAATTTGTGGCA 122
61 AATTAGGAAAAAGAAAAAGAAAAAGAGAGAGAGAGAAATTCACGCCAAATTTGTGGCA 120
123 TAGATTAT 182
121 TAGATTAT 180
183 AAGCCCTGTTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGAGAAATGTTAACGGAG 242

181 AAGCCCTGTTGTGTCCACCATCTCCAAAGAGGTTTACCTGCAGGAAATGTTTAAACGGAG 240
243 GCTGCTTCCCTGGGCAACAGGAGCCACCTGGCGCAGGACGCTTTTTCAGAAAGAGACGC 302
241 GCTGCTTCCCTGGGCAACAGGAGCCACCTGGGCGAG----- 277
303 CTTTTCAGGAAGAGAGCGCTTTTTCAGGAAGAGAGAAAGTGCAGTGAAGAGAAAGTCAAC 362
278 -----GAGAAAGTGCAGCTGAAGAGAAAGTCAAC 306
363 TTTACTGAGGGGAGTCTCCATTTATCATTTGGGACCATCATTTGGGACGAGAAATCTTCATCTC 422
307 TTTACTGAGGGGAGTCTCCATTTATCATTTGGGACCATCATTTGGGACGAGAAATCTTCATCTC 366
423 TCTTAAGGCGTGTCTCCAGAACACGCGGCGAGGCTGGGCAATGCTCTGACCAATCTGGAGCGT 482
367 TCTTAAGGCGTGTCTCCAGAACACGCGGCGAGGCTGGGCAATGCTCTGACCAATCTGGAGCGT 426
483 GTGTGGGGTCTGTCTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAAACAATATAAA 542
427 GTGTGGGGTCTGTCTCACTATTTGGAGCTTTGTCTTATGCTGAATTTGGGAAACAATATAAA 486
543 GAAATCTGGAGGTCATTTACACATATATTTTGAAGTCTTTTGGTCCATTTACAGCTTTTGT 602
487 GAAATCTGGAGGTCATTTACACATATATTTTGAAGTCTTTTGGTCCATTTACAGCTTTTGT 546
603 ACAGTCTGGGTGGAACTCTCTCATATACGCCCTGAGCTACTGCTGTGATATCCCTGGC 662
547 ACAGTCTGGGTGGAACTCTCTCATATACGCCCTGAGCTACTGCTGTGATATCCCTGGC 606
663 ATTTGGAGCTACATTTCTGGAACCATTTTATTTCAATGTGAATCCCTGAACTTGGCAT 722
607 ATTTGGAGCTACATTTCTGGAACCATTTTATTTCAATGTGAATCCCTGAACTTGGCAT 666
723 CAAGCTCATTTACAGCTGTGGGCATAACTGTAGTGTGGTCTCTAAATAGCATGAGTGTGAG 782
667 CAAGCTCATTTACAGCTGTGGGCATAACTGTAGTGTGGTCTCTAAATAGCATGAGTGTGAG 726
783 CTGAGGCGCGGATCCAGATTTTCTTAACCTTTTCAAGCTCCACAGCAATTTCTGATAAT 842
727 CTGAGGCGCGGATCCAGATTTTCTTAACCTTTTCAAGCTCCACAGCAATTTCTGATAAT 786
843 TATAGTCTCTGGAGTTATGAGCTTAATTAAGGTCAAAAGCAGAACTTTTAAAGACGCTTT 902
787 TATAGTCTCTGGAGTTATGAGCTTAATTAAGGTCAAAAGCAGAACTTTTAAAGACGCTTT 846
903 TTCAGGAAGAGATTCAGAGTATTCAGCGGTTGCCACTGGCTTTTATATGGAATGTATGC 962
847 TTCAGGAAGAGATTCAGAGTATTCAGCGGTTGCCACTGGCTTTTATATGGAATGTATGC 906
963 ATATGCTGGCTGGTTTACCTCAACTTTGTACTGAAGAGTAGAAACCCCTGAAAAAAC 1022
907 ATATGCTGGCTGGTTTACCTCAACTTTGTACTGAAGAGTAGAAACCCCTGAAAAAAC 966
1023 CATTCCTCTGCAATPATGTATTCATGCGCAATTTGTCACCATTTGGCTATGCTGACAAA 1082
967 CATTCCTCTGCAATPATGTATTCATGCGCAATTTGTCACCATTTGGCTATGCTGACAAA 1026
1083 TGTGGCTCTACTTTAGCAGCAATTAATGCTGAGAGCTGCTGCTTTTCAAAATGCAAGTGGCAGT 1142
1027 TGTGGCTCTACTTTAGCAGCAATTAATGCTGAGAGCTGCTGCTTTTCAAAATGCAAGTGGCAGT 1086
1143 GACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATATTAGCAGTTCGGATCTTTGTTGCCCT 1202
1087 GACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATATTAGCAGTTCGGATCTTTGTTGCCCT 1146
1203 CTCCTGCTTTGGCTCCATGAAAGGCTGCTGCTGCTGCTCCAGGTTATTTCTATGTTGC 1262
1147 CTCCTGCTTTGGCTCCATGAAAGGCTGCTGCTGCTGCTGCTCCAGGTTATTTCTATGTTGC 1206
1263 GTCTCAGAGAGGCTACCTTTCCAGAAATCCTCTCCATGATTCATGTCGCAAGACACTCC 1322

; APPLICANT: Fling, Steven P.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Indrias, Carol Joseph
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Elliot, Mark
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Kalos, Michael D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; FILE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
 ; CURRENT APPLICATION NUMBER: US/09/738,973
 ; CURRENT FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 587
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 440
 ; LENGTH: 2239
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-738-973-440

Query Match 100.0%; Score 2239; DB 10; Length 2239;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGAGGTTGAAGTGCAGAGATCATGCCAGCCCTGGGTGACGTGAGACTCTGCTCAAAAC	60
DB	1	GGAGGTTGAAGTGCAGAGATCATGCCAGCCCTGGGTGACGTGAGACTCTGCTCAAAAC	60
QY	61	AGATTAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG	120
DB	61	AGATTAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG	120
QY	121	CATAGATTTTATCATATCTCGATTTTGGATTTCTTTGTTTCTCATCATGATTTCA	180
DB	121	CATAGATTTTATCATATCTCGATTTTGGATTTCTTTGTTTCTCATCATGATTTCA	180
QY	181	GGAAGCCCTGTTGTCACCATCTCCAAAGAGAGTTACCTGCAGGGAATGTTAAAGGG	240
DB	181	GGAAGCCCTGTTGTCACCATCTCCAAAGAGAGTTACCTGCAGGGAATGTTAAAGGG	240
QY	241	AGGCTGCCCTCCCTGGGCAACAGGAGCCACCTGGGAGGAGGCTTTTTCAGGAAGAGAC	300
DB	241	AGGCTGCCCTCCCTGGGCAACAGGAGCCACCTGGGAGGAGGCTTTTTCAGGAAGAGAC	300
QY	301	GCCTTTTCAGGAAGAGAGCCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGGAAAGTC	360
DB	301	GCCTTTTCAGGAAGAGAGCCCTTTTCAGGAAGAGAGAAAGTGCAGCTGAAGGAAAGTC	360
QY	361	ACTTTACTGAGGGAGTCTCCATTTATCATTTGACCATCATTTGAGGAGAAATCTTCATC	420
DB	361	ACTTTACTGAGGGAGTCTCCATTTATCATTTGACCATCATTTGAGGAGAAATCTTCATC	420
QY	421	TCCTCTAAGGGGCTGCTCCGAAACAGGAGCCAGCTGGGAGTCTGTCACCATCTGGAGC	480
DB	421	TCCTCTAAGGGGCTGCTCCGAAACAGGAGCCAGCTGGGAGTCTGTCACCATCTGGAGC	480
QY	481	GTGTGTGGGGTCCCTGCTCACTATTGGAGCTTTGCTTTATCTGCTGAATGGGAACAATA	540
DB	481	GTGTGTGGGGTCCCTGCTCACTATTGGAGCTTTGCTTTATCTGCTGAATGGGAACAATA	540
QY	541	AAGAAATCGAGGTCATTTACATATATTTTGGAGTCTTTGGTCCATTTACAGGCTTTT	600
DB	541	AAGAAATCGAGGTCATTTACATATATTTTGGAGTCTTTGGTCCATTTACAGGCTTTT	600
QY	601	GTACAGTCTGGGGTGAACCTCCATATACGCGCTGACGCTACTGCTGTATATCCCTG	660
DB	601	GTACAGTCTGGGGTGAACCTCCATATACGCGCTGACGCTACTGCTGTATATCCCTG	660
QY	661	GCATTTGGAGCTACATTTCTGGAACCATTTTATTTCAATGTAATCTTCACTTGGG	720
DB	661	GCATTTGGAGCTACATTTCTGGAACCATTTTATTTCAATGTAATCTTCACTTGGG	720

QY	721	ATCAAGCTCAATACAGCTGTGGGATTAACCTAGTGTGCTTAAATAGCATGAGTGC	780
DB	721	ATCAAGCTCAATACAGCTGTGGGATTAACCTAGTGTGCTTAAATAGCATGAGTGC	780
QY	781	AGCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATTCGATA	840
DB	781	AGCTGGAGCGCCGGATCCAGATTTCTTAACCTTTTGAAGCTCACAGCAATTCGATA	840
QY	841	ATTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCACACGAGAACTTTAAAGACGC	900
DB	841	ATTATAGTCCCTGGAGTTATGAGCTAATTAAGGTCACACGAGAACTTTAAAGACGC	900
QY	901	TTTTTCAGGAAGAGATTCAGATTTACCGGTTGCCACTGGCTTTTATTTATGGAATGAT	960
DB	901	TTTTTCAGGAAGAGATTCAGATTTACCGGTTGCCACTGGCTTTTATTTATGGAATGAT	960
QY	961	GCATATGCTGGCTGGTTTACCTCAACTTTGTTTACTGAAAGAGTACAAAACCTGAAAA	1020
DB	961	GCATATGCTGGCTGGTTTACCTCAACTTTGTTTACTGAAAGAGTACAAAACCTGAAAA	1020
QY	1021	ACCATTCCCTTCCAATATGATATCCATGGCCATTGTCACCATTTGGCTATGCTGACA	1080
DB	1021	ACCATTCCCTTCCAATATGATATCCATGGCCATTGTCACCATTTGGCTATGCTGACA	1080
QY	1081	AATGTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATCCAGTGCA	1140
DB	1081	AATGTGGCTTACTTTACGACCATTAATGCTGAGGAGCTGCTGCTTCAAAATCCAGTGCA	1140
QY	1141	GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTC	1200
DB	1141	GTGACCTTTTCTGAGCGGCTACTGGGAAATTTCTCATTAGCAGTTCCGATCTTTGTC	1200
QY	1201	CTCTCTGCTTTGGCTCCATGAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1260
DB	1201	CTCTCTGCTTTGGCTCCATGAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1260
QY	1261	GGCTCTGAGAGGGTCACTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACT	1320
DB	1261	GGCTCTGAGAGGGTCACTCCAGAAATCCTCTCCATGATTCATGTCGCAAGCACACT	1320
QY	1321	CCTCTACAGCTGTTTATTTGTTTGGACCTTTGACCAATGATAATGCTCTCTCTGAGAC	1380
DB	1321	CCTCTACAGCTGTTTATTTGTTTGGACCTTTGACCAATGATAATGCTCTCTCTGAGAC	1380
QY	1381	CTGACAGCTCTTTGAATTTCCCTCAGTTTCCAGAGTGGCTTTTATTTGGGCTGGAGTT	1440
DB	1381	CTGACAGCTCTTTGAATTTCCCTCAGTTTCCAGAGTGGCTTTTATTTGGGCTGGAGTT	1440
QY	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATGCTCTTCAAGGTGCCA	1500
DB	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGATGCTCTTCAAGGTGCCA	1500
QY	1501	CTGTTTATCCAGCTTTTCTTCCATGCTCTCTCATGCTTCCCTCTCTAT	1560
DB	1501	CTGTTTATCCAGCTTTTCTTCCATGCTCTCTCATGCTTCCCTCTCTAT	1560
QY	1561	TCGGACCATTTAGTACAGGATTTGGCTTCGTCATCCTGCTGAGTCCCTGCTGAT	1620
DB	1561	TCGGACCATTTAGTACAGGATTTGGCTTCGTCATCCTGCTGAGTCCCTGCTGAT	1620
QY	1621	TATCTCTTTTATATATGGGACAAAGAACCCAGGTGGTGTAGATAATGTCAGAGAAATA	1680
DB	1621	TATCTCTTTTATATATGGGACAAAGAACCCAGGTGGTGTAGATAATGTCAGAGAAATA	1680
QY	1681	ACCAGACATTTACAATATGAGTGGAGTTGACACAGAAAGATTAAGTATGAACATA	1740
DB	1681	ACCAGACATTTACAATATGAGTGGAGTTGACACAGAAAGATTAAGTATGAACATA	1740
QY	1741	GGACTTGGAGCTTTGGCAATCTGCCCAAGGGGAGACACAAATAGGATTTTACTTCAT	1800
DB	1741	GGACTTGGAGCTTTGGCAATCTGCCCAAGGGGAGACACAAATAGGATTTTACTTCAT	1800

QY	1321	CCTCTCAGAGCTGTTATTTGTTTTCACCCCTTGACAAATGATAATGCTCTTCTCTCGAGAC	1380
Db	1321		
QY	1381	CCTCTACAGAGCTGTTATTTGTTTTCACCCCTTGACAAATGATAATGCTCTTCTCTCGAGAC	1380
Db	1381		
QY	1381	CCTCTACAGAGCTGTTATTTGTTTTCACCCCTTGACAAATGATAATGCTCTTCTCTCGAGAC	1440
Db	1381		
QY	1441	GCTGGGCTGATTTATCTTCGATACAAATGCCAGATATGCAATCGCTCTTCAAGGTGCCA	1500
Db	1441		
QY	1501	CTGTTTCATCCAGCTGTTGTTTTCCTTCACATGCCCTTCCTTCATGGTTGCCCTTCCTCTAT	1560
Db	1501		
QY	1561	TCGGAGCCATTTAGTACAGAGGATTCGCTTCGTCATCTCTGACTGGAGTCCCTGCGTAT	1620
Db	1561		
QY	1621	TATCTCTTTTATATATGGGACAAAGAACCCAGGTGGTTTAGAATTAATGTCAGAGAAATA	1680
Db	1621		
QY	1681	ACCAGAACATTAACAATAATGCTGGAAGTTGTACCAGAGAGAGATAAGTTATGAACATAAT	1740
Db	1681		
QY	1741	GGACTTCAGATCTTGGAATTTACAACTTCGCCAAGGGAGACACAAATAGGGATTTTACTTCAT	1800
Db	1741		
QY	1801	TTCTGAAAGCTTAGAGATTTACAACTTTGGTGATTAACAAAGAGAGTCAAGTTATTTTAA	1860
Db	1801		
QY	1861	TTCTGAAAGCTTAGAGATTTACAACTTTGGTGATTAACAAAGAGAGTCAAGTTATTTTAA	1920
Db	1861		
QY	1921	TTATAGAAAGTGAATATGCACTTATCTATGAGTGGCACAATTTCTGAGTCTCTGATACC	1980
Db	1921		
QY	1981	TACCTATTTGGGTTAGGAGAAAGACTAGACAATTTACTATGCTGCTATCTCTACAACT	2040
Db	1981		
QY	2041	ATGTTAGCAGCGCAAGAACCTTCAAAATTTGAAGACTGAGATTTTCTGTATATATGGTT	2100
Db	2041		
QY	2101	TTGTAAGAGTGGTTTACACACTACAGATGCTATACTGTGAAAGTGTTCATATCTG	2160
Db	2101		
QY	2161	AAAAAGCATACATCATGATTTATGGCAAGAGGAGAGAGTAGAGCTGTCTTAAAT	2220
Db	2161		
QY	2221	TATTAAGAGTGGTTTACACACTACAGATGCTATACTGTGAAAGTGTTCATATCTG	2280
Db	2221		

QY	241	AGGCTGCCCTCCCTGGCAAGAGGACCACTGGCGAGGACGCTTTTTCAGGAAGAGAC	300
Db	241		
QY	301	GCCTTTTCAGGAAGAGAGAGGCTTTTCAGGAAGAGAGAGAGAGAGAGAGAGAGAGC	360
Db	301		
QY	361	ACTTTTACTGAGGGAGTCTCCATTTATCATTTGGCAGGAGGATCTTTCATC	420
Db	361		
QY	421	TCCTCTAAGGGCTGCTCCAGAACAGCGGACGCTGGGATGTCCTGACCATCTGGAGC	480
Db	421		
QY	481	GTGTGGGGTCTCTGACATTTTGGAGCTTTGCTTATGCTGAATTTGGGAACAACTATA	540
Db	481		
QY	541	AAGAAATCTGGAGTCAATACACATATATTTTGGAAAGTCTTTGGTCCATACCACTTTT	600
Db	541		
QY	601	GTAGAGTCTGGTGGGACCTCTCATATACGCCCTGCAGCTACTGCTGTATATCCCTG	660
Db	601		
QY	661	GCATTTGGAGCTACATCTTGGAAACATTTTATTAATGTGAATCCCTGAACTTGGG	720
Db	661		
QY	721	ATCAAGCTCATACAGCTGTGGGATTAACCTTTTAAAGTGTAGTGTCTTAATAGCATGTC	780
Db	721		
QY	781	AGCTGGAGCGCCGATTCAGATTTTCTTAACCTTTTGAAGCTCACAGCAATTTCTGATA	840
Db	781		
QY	841	ATTATAGTCCCTGGAGTTATCAGCTAATTAAGGTCAACGAGCACTTTTAAAGAGCC	900
Db	841		
QY	901	TTTTTCAGGAAGAGATTTCAAGTATACGGGTTGCGGCTTTTATATATGGAATGTAT	960
Db	901		
QY	961	GCATATGCTGGCTGCTTTTACCTCAACTTTGTTACTGAAGAGTGAAGAACCCCTGAAAA	1020
Db	961		
QY	1021	ACCATTCCCTTGCATATGATATATATGCAATGTCACCATTTGCTATGCTGAC	1080
Db	1021		
QY	1081	AATGTGGCTACTTTACGAGCAATTAATGCTGAGGAGTGTGCTTTCAAAATGCAAGTGCA	1140
Db	1081		
QY	1141	GTGACCTTTTCTGAGGGCTACTGGGAAATTTCTCATTTAGCAGTTCCGATCTTTGTTGCC	1200
Db	1141		
QY	1201	CTCTCTGCTTTGGCTTCCATGACGGTGTGTTTGTCTGCTCCAGGTTATCTATGTT	1260
Db	1201		
QY	1261	CGGCTCGAGAGGCTCACCTTCAGAAATCCCTCCATGATTCATGTCGCGAAGCACACT	1320
Db	1261		

RESULT 2
US-09-738-973-440
; Sequence 440, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.

GenCore version 5.1.3

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 16, 2003, 06:28:33 ; Search time 232 Seconds
(without alignments)
9694.889 Million cell updates/sec

Title: US-09-667-170A-440

Perfect score: 2239

Sequence: 1 ggaggttgagtgagcagag.....ttattataaaaaaaaaa 2239

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2239	100.0	2239	9	US-09-854-133-440
2	2239	100.0	2239	10	US-09-738-973-440
3	2080.8	92.9	5981	9	US-09-854-133-441
4	2080.8	92.9	5981	10	US-09-738-973-441
5	1900.2	84.9	3144	9	US-10-163-866-31
6	1888.2	84.3	2482	9	US-10-163-866-29
7	1540.8	68.8	1861	9	US-10-163-866-30
8	1454.4	65.0	1542	9	US-10-163-866-33
9	1431.8	63.9	1528	9	US-10-163-866-52
10	1373.4	61.3	2000	9	US-10-163-866-34
11	1047.4	46.8	1268	9	US-10-163-866-53
12	449	20.1	520	9	US-10-163-866-32
13	379.2	16.9	1581	9	US-10-163-866-10
14	379.2	16.9	6295	9	US-10-163-866-8
15	379.2	16.9	6296	9	US-10-163-866-9
16	366.4	16.4	1582	9	US-10-163-866-12
17	365.4	16.3	1609	9	US-10-163-866-5
18	365.4	16.3	1621	9	US-10-163-866-7
19	365.4	16.3	4559	10	US-09-919-172-61

20	365.4	16.3	4559	12	US-10-044-090-555
21	364.8	16.3	2098	9	US-10-163-866-11
22	362.2	16.2	1541	9	US-10-163-866-51
23	362.2	16.2	1559	9	US-10-163-866-3
24	362.2	16.2	1593	9	US-10-163-866-1
25	362.2	16.2	4539	9	US-10-163-866-2
26	343.8	15.4	3728	9	US-10-163-866-14
27	343.8	15.4	4237	9	US-10-163-866-15
28	343.8	15.4	4237	9	US-10-163-866-20
29	342.2	15.3	1657	9	US-10-163-866-16
30	337.2	15.1	3717	9	US-10-163-866-13
31	319.6	14.3	2529	9	US-09-815-923-15
32	315.6	14.1	1595	9	US-10-163-866-27
33	315.6	14.1	1918	9	US-10-163-866-26
34	311.6	13.9	1962	9	US-10-163-866-28
35	301	13.4	1476	9	US-10-163-866-24
36	301	13.4	1654	9	US-10-163-866-22
37	301	13.4	1664	9	US-10-163-866-25
38	299.4	13.4	1814	9	US-10-163-866-23
39	297.8	13.3	1651	9	US-10-163-866-21
40	281	12.6	337	9	US-09-854-133-442
41	281	12.6	337	10	US-09-738-973-442
42	238.4	10.6	3984	9	US-09-954-531-151
43	238.4	10.6	3984	9	US-10-163-866-6
44	230	10.3	572	9	US-10-046-935-1307
45	230	10.3	572	9	US-09-878-178-1307

ALIGNMENTS

RESULT 1

US-09-854-133-440
 ; Sequence 440, Application US/09854133
 ; Publication No. US20020183499A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Secrist, Heather
 ; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.475C10
 ; CURRENT APPLICATION NUMBER: US/09/854,133
 ; CURRENT FILING DATE: 2001-05-11
 ; NUMBER OF SEQ ID NOS: 735
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 440
 ; LENGTH: 2239
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-854-133-440

Query Match	100.0%	Score	2239;	DB	9;	Length	2239;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	2239;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	GGAGGTGAGTACGACAGATCATGCGCTGGTGACAGTCTGTCTCAAC	60				
Db	1	GGAGGTGAGTACGACAGATCATGCGCTGGTGACAGTCTGTCTCAAC	60				
QY	61	AGATTAAAGGAAAAAGAAAGAAAGAGAGAGAGAAATCCAGGCCAATTTGG	120				
Db	61	AGATTAAAGGAAAAAGAAAGAAAGAGAGAGAGAAATCCAGGCCAATTTGG	120				
QY	121	CATACATTTTATCATATTTCTGGATTTTCTTTTCTCATCTGATTTCA	180				
Db	121	CATACATTTTATCATATTTCTGGATTTTCTTTTCTCATCTGATTTCA	180				
QY	181	GGAAAGCCTGTGTCTCCACCATCTCCAAAGGAGGTTCCTCGAGGAATGTACGG	240				
Db	181	GGAAAGCCTGTGTCTCCACCATCTCCAAAGGAGGTTCCTCGAGGAATGTACGG	240				

; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match 1.9%; Score 42.8; DB 4; Length 72604;
Best Local Similarity 69.4%; Pred. No. 0.34;
Matches 75; Conservative 0; Mismatches 27; Indels 6; Gaps 1;
QY 1 GGAGGTTGAAGTCAGCAGAGATCATGCCA-----GCCTGGGTGACAGTGTGACTCTGTC 54
Db 45080 GGAGGTTGAGTCAGCAGCCGAGATGGTCCCATTCGACTCCAGCCTGAGTGACAGAGCAAGAC 45139
QY 55 TCAACAGAGATTAAAGGAAAAAGAAAAAGAAAAAGAGAGAGAGGAA 102
Db 45140 TCCATCTCAATAAGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAAA 45187

Search completed: April 16, 2003, 09:33:58
Job time : 1056 secs

[illegible]

LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031

OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57